DOCUMENT RESUME

ED 475 764 TM 034 885

TITLE	North East Association for Institutional Research Annual Conference Proceedings (29th, Annapolis, Maryland, November 16-19, 2002).
INSTITUTION	North East Association for Institutional Research.
PUB DATE	2002-11-00
NOTE	149p.
PUB TYPE	Collected Works - Proceedings (021)
EDRS PRICE	EDRS Price MF01/PC06 Plus Postage.
DESCRIPTORS	Computer Uses in Education; Conferences; *Educational Assessment; Educational Research; *Higher Education; *Institutional Research

ABSTRACT

This proceedings contains papers from the 2002 annual conference of the Northeast Association for Institutional Research, a meeting devoted to assessment in the 21st century and the challenges that face institutional research. The papers are: (1) "Putting Community College Enrollment Trends in Perspective by the Use of Census Data and Market Measurement Techniques" (Karl Boughan and Ophelia Robinson); (2) "Linking Assessment Planning and Mission Review: One University's Experience" (Ellen Boylan-Fick and Barbara R. Sadowski); (3) "An Assessment of Standardized Accuplacer Placement Scores for College English in the Connecticut Community-Technical College System" (Corby A. Coperthwaite and William F. Ritchie); (4) "Enhancing Outcomes Assessment by Discovering Alumnae/i Success Strategies" (Anne Marie Delaney); (5) "Determinants of Student Dropout in Critical Periods: Cohort Differences at a Virtual University" (Tae Young Han and Mitchell S. Nesler); (6) "Linking Internal Transfer Patterns to College Student Experiences: A Case Study" (Jean O. Marriott and Catherine J. Alvord); (7) "An Analysis of the Retention of First-Time Full-Time Freshmen at a Public Urban University" (Kevin B. Murphy); (8) "The Impact of Contact Type on Web Survey Response Rates" (Stephen R. Porter and Michael E. Whitcomb); (9) "Statistical Methods for Predicting Yield: A Comparison of the Accuracy of Logistic Regression, Decision Tree, Neural Network and Boosted Logistic Regression" (Richard J. Reeves and Martin T. Wells); (10) "Paper vs. Web: The Differential Impact on Responses of Men and Women" (Heather S. Roscoe and Dawn Geronimo Terkal); (11) "The Frequencies of Student Online Activity as Predictors of Course Grade" (Charles Secolsky and Cliff L. Wood); (12) "Using Grid-Group Theory To Understand Students and Institutions" (Carol Trosset); (13) "Assessment and Assisting the College President Steer the Ship: An Analytic Comparison of Dashboard Indicators, the Balanced Scorecard, Performance Measures, and Six Sigma in the College and University Setting" (Gail Wisan); and (14) "Closing the Assessment Loop: Applying Results of a Primary Trait Analysis To Improve Educational Outcomes" (David W. Wright, Elizabeth A. Robinson, and Marsha V. Krotseng). Each paper contains references. (Contains 10 figures, 10 charts, 2 diagrams, and 42 tables.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 475 764

North East Association for Institutional Research

29th Annual Conference Proceedings

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

B. A. Simpson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

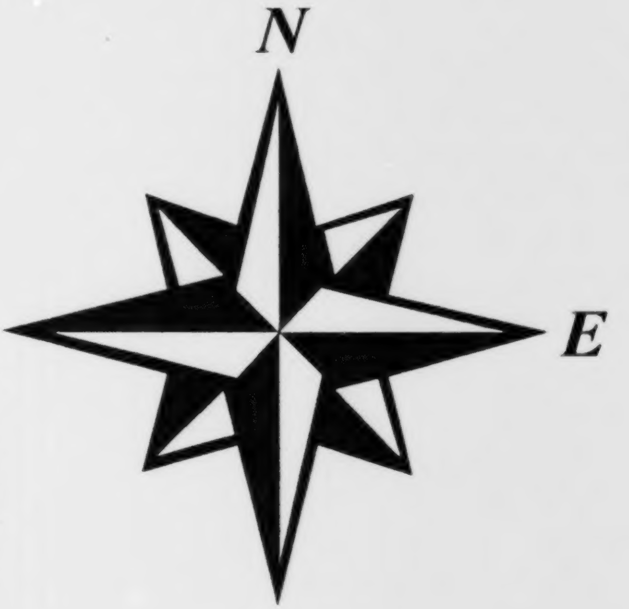
1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.



Loews Hotel
Annapolis, MD
November 16-19, 2002

TM034885

"Assessment in the 21st Century: Challenges for IR"

BEST COPY AVAILABLE

Dear NEAIR Friends and Colleagues,

The 2002 annual conference of the Northeast Association for Institutional Research (NEAIR) was held from November 16th to 19th at the Loews Hotel in Annapolis, Maryland. The theme for the conference was "*Assessment in the 21st Century: Challenges for IR.*" A record-breaking crowd of nearly 280 individuals attended the conference.

The conference program included 15 pre-conference workshops, 3 plenary sessions, 5 panel sessions, 35 paper presentations and the variety of poster sessions, table topics, vendor showcases, and special interest group meetings. Keynote addresses were delivered by James Anderson, Vice Provost for Undergraduate Affairs at North Carolina State, who spoke on *Generating Success Models that Integrate the Assessment of Student Learning and Engagement, Effective Teaching and Diversity* and Patricia Haeuser, Director of Planning and IR at Northern Arizona University, who addressed *Assessment for Institutional Improvement*. Dawn Terkla, Executive Director of IR at Tufts and past NEAIR and AIR President, led a panel on *Professional Development Paths* which included Marian Pagano, Associate Provost for Planning and Institutional Research at Columbia University, Elizabeth Sibolski, Executive Associate Director of the Middle States Association Commission on Higher Education, Stephen Thorpe, Vice President for Academic Affairs at Neumann College, and James Tschechtelin, recently retired President of Baltimore City Community College. All four of these individuals worked in institutional research earlier in their careers and all four had been active in our professional association at one time or another.

Special events included a banquet with an Annapolis flair highlighted by performances by the U.S. Naval Academy's Silent Drill Team and Gospel Choir. Despite soggy weather, attendees also enjoyed local entertainment, trips to the U.S. Naval Academy and tours of the historic area. A number of activities were held that were designed to be of particular interest to newcomers including a newcomers' pre-conference workshop, the career paths panel and a welcoming reception which focused on our mentorship program.

A true highlight of the conference was the granting of NEAIR's Distinguished Service Award to Dawn Terkla – only our third such recipient. Dawn received a well deserved standing ovation from the conference attendees in appreciation for the many outstanding contributions she has made to our profession and organization over the years!

By virtue of a vote taken at the annual business meeting held in conjunction with the conference, NEAIR became the first regional organization to offer a formal endorsement of the AIR Code of Ethics, bringing the development of the Code nearly full-circle as many of the AIR members who were instrumental in developing it hailed from the northeast.

Of course, planning and running an event of the magnitude of the NEAIR annual conference would be impossible without the tireless efforts of many dedicated volunteers. As an organization we owe a great debt of gratitude to the planning team for the Annapolis conference. Program Chair Michelle Appel and Associate Program Chair Tracy Hunt-White did a marvelous job pulling together an outstanding and engaging program. Local Arrangements Chair Ella Smith was extraordinary in creating a welcoming atmosphere for the conference with a particularly Annapolis flair. The efforts of Alan Harmon, from the U.S. Naval Academy, Martha Gray and the members of the Newcomers Committee, Evaluation Chair Mindy Wang, Publications Chair Marianthi Zikopoulos, Web Designer Meihua Zhai, NEAIR Treasurer Corby Coperthwaite and the members and officers of MdAIR all helped make the conference one that will be long remembered. The advice and support offered by the 2001-2002 NEAIR Steering Committee and, in particular, the 2001 conference planning team led by Anne Marie Delaney, Kelli Armstrong and Bea Frain proved invaluable. Of course, it goes without saying that none of this would be possible without the exceptional assistance and encouragement offered by our Membership Secretary, Beth Simpson.

I encourage you to read closely the documents included in this publication of our 2002 Conference Proceedings. I believe you will find that they highlight some of the best work that our field has to offer. I would welcome any comments you may have about either the 2002 conference, in general, or this publication, in particular. I look forward to seeing you at many NEAIR events in the years to come.

Sincerely,

Jim Trainer

James F. Trainer
Villanova University
NEAIR President, 2001-2002

2001 - 2002 Steering Committee

Officers

President
Jim Trainer
Villanova University

President Elect
Becky Brodigan
Middlebury College

Past President
Anne Marie Delaney
Babson College

Secretary
Sarah Parrott
Brandeis University

Treasurer
Corby Coperthwaite
Connecticut Community-Technical
College System

Members at Large

Fred Cohen
New York University

Martha Gray
Ithaca College

Heather Kelly Isaacs
University of Delaware

John Pryor
Dartmouth College

Stephen Porter
Wesleyan University

Membership Secretary
Beth Simpson
HEDS Consortium

2002 Conference Chairs

Program Chair
Michelle Appel
University of Maryland, College Park

Associate Program Chair
Tracy Hunt White
Education Statistics Services Institute

Local Arrangements Chair
Ella Smith
Loyola College

Publications Chair
Marianthi Zikopoulos
New School University

TABLE OF CONTENTS

Putting Community College Enrollment Trends in Perspective by the Use of Census Data and Market Measurement Techniques 1
Karl Boughan

Linking Assessment Planning and Mission Review: One University's Experience 11
Ellen Boylan-Fick and Barbara R. Sadowski

An Assessment of Standardized Accuplacer Placement Scores for College English in the Connecticut Community-Technical College System 23
Corby A. Coperthwaite and William F. Ritchie

Enhasing Outcomes Assessment by Discovering Alumnae/I Success Strategies 33
Anne Marie Delaney

Determinants of Student Dropout in Critical Periods: Cohort Differences at a Virtual University 45
Tae Young Han and Mitchell S. Nesler

Linking Internal Transfer Patterns to College Student Experiences: A Case Study 55
Jean O. Marriott and Catherine J. Alvord

An Analysis of the Retention of First-time Full-time Freshmen at a Public Urban University 67
Kevin B. Murphy

The Impact of Contact Type on Web Survey Response Rates*** 79
Stephen R. Porter and Michael E. Whitcomb

Statistical Methods for Predicting Yield: A Comparison of the Accuracy of Logistic Regression, Decision Tree, Neural Network, and Boosted Logistic Regression 95
Richard J. Reeves and Martin T. Wells

Paper vs. Web: The Differential Impact on Responses of Men and Women 105
Heather S. Roscoe and Dawn Geronimo Terkla

The Frequencies of Student Online Activity as Predictors of Course Grade 115
Charles Secolsky and Cliff L. Wood

Using Grid-Group Theory to Understand Students and Institutions 125
Carol Trosset

Assessment and Assisting the College President Steer the Ship: An Analytic Comparison of Dashboard Indicators, The Balanced Scorecard, Performance Measures, and Six Sigma In the College and University Setting 131
Gail Wisan

Closing the Assessment Loop: Applying Results of a Primary Trait Analysis to Improve Educational Outcomes 141
David W. Wright, Elizabeth A. Robinson, and Marsha V. Krotseng

2002 Conference Program 147

*****This paper received the 2002 Best Paper Award**

PUTTING COMMUNITY COLLEGE ENROLLMENT TRENDS IN PERSPECTIVE BY THE USE OF CENSUS DATA AND MARKET MEASUREMENT TECHNIQUES

Karl Boughan
Director of Institution Research

Ophelia Robinson
Research Analyst

Prince George's Community College

Introduction

There is no *text* without *con*text, and the proper contexts for reading community college enrollment data are the geo-demography and market dynamics of the designated service area of the school. How, for instance, can one get an adequate sense of the ethnic diversity of one's student body without comparing its cultural breakdown with that of the surrounding territory, or correctly ascertain whether an enrollment group's growth is a product of larger demographic trends or the result of successful recruitment campaigning? And how assured can the institutional researcher be in his site recommendations for a new extension center or campus if he has neglected measuring candidate locales for evidence of enrollment under-servicing?

We find it strange, then, that the published literature contains few examples of systematic efforts at integrating community college enrollment and geo-demographic analysis. There are student body fact books aplenty and environmental scans by the score, but few seem inclined to put the two knowledge bases together into a truly informative package. Some of the studies representing this rare strain are those of Hecht (1991), Heacock and Jenkins (1993) and Baldwin (1998), which demonstrate the utility of Census data and market analysis for individual community colleges; the National Center for Higher Educational Management Systems (1998) and the Virginia State Department of Community Colleges (2000), which illustrate the use of demographic penetration rates to identify the market circumstances of entire state systems; and Newmyer and Mangham (1992) and Boughan (1993), which center on special applications of geo-demographic educational market research – campus site location and community outreach targeting, respectively.

In this paper we will present the methodology and findings of a geo-demographic market study of a single school: Prince George's Community College, a large, suburban two-year institution serving Prince George's County, Maryland, on the eastern border of Washington, D.C. Our research plan involved a systematic scanning for relationships between County demographic and PGCC credit enrollment shifts over the decade of the 1990s.

To do this, we employed two sets of student record data representing all unduplicated enrollees during school fiscal years 1990-1991 and 1999-2000, and two County, tract-level data sets derived from the U.S. Census of 1990 and of 2000. County population analysis was restricted to those qualifying as potential community college students – *adult residents* (those aged 18 years or older) – and only *county residents* were included in the study's student body data sets. Three key enrollment management variables, Race/Ethnicity, Age and Income, were selected for treatment and basic county/student relational patterns were captured by a common set of distributions and market measures – per year absolute numbers, percentages and draw rates, percent changes over time, and disproportion indices. Then, as a logical follow-up we divided the county into demographically meaningful regions by aggregating Census tracts, and performed a truncated version of the countywide analysis in order to identify areas under-served by PGCC. The last portion of the paper presents data bearing on the success of a new extension center established largely as a result of this type of analysis.

Methodological Considerations

Our study's design was informed by a particular conception of the layered structure of two-year postsecondary markets. At the top of an inverted pyramid is the *total service area population*, followed by the *standing market* for a community college, all potential two-year students (adult residents). The residuum, those *under* 18 years old, influence the *future* standing market according to their numbers and the timing of their entrance into the standing market. The *practical market*, which comes next, is a subset of the standing market – basically its *non-college graduate* part, those with a the most evident need of the degrees and occupational certificates provided by a community college. Since Census data exist on college degree-holding (educational attainment), it is a relatively easy task to derive practical market data sets, but there is an empirical problem related to using such in a community college setting since a discernable minority of students at two-year school are in fact college graduates seeking career change and other non-degree-related objectives. The offspring of the practical market is the *ready market*, potential community college students actually planning on, or strongly desiring to, attend. Unfortunately, this group is usually beyond institutional research consideration because it requires survey data for definition. Finally, at the bottom of the pyramid we find the *actualized* or *active market*, currently enrolled community college students, and its most interesting component – the current *customer base* of the researcher's own college.

While all of these markets are worth researching, given the difficulties of operationalizing the intervening levels and the fact that this study is our first attempt at systematic market analysis, we will concentrate almost entirely upon the relationship between the top and bottom of the full market spectrum – the county's higher education *standing market* and PGCC's *customer base.* To that end we will utilize the main marketing measure of that relationship, known as market *penetration* or *draw rate* (sometimes called the *participation rate* in the institutional research literature). This is simply the customer base (here, all county-resident PGCC credit enrollees in a fiscal year) divided by the standing market (all county adult residents), expressed as a percentage. Lastly, it should be underlined that *draw rate* is *not* synonymous with *market share*, a measure common in institutional research. Market share refers only to the portion of the *active market* which falls to a school rather than to rival institutions.

On the data side, the main set-up work was achieving alignment of Census data with student record data, market category by category. Age group parallelism was not a problem since PGCC student ages can be sorted by any breakdown scheme desired, and for both 1990 and 2000 Census data collections, the Bureau provides sufficiently finely cut age counts to arrange categories as needed. Census age data, in fact most of the Census data used in this study, was accessed by means of the Bureau's Internet site *American FactFinder*. The relevant files in this case were 1990 Summary File Tape 1 (STP1) and 2000 Summary File 1 (SF1).

Achieving Race/Ethnicity parallelism, however, required a bit more effort. PGCC uses a single registration question-derived five point variable of Race/Ethnicity, directly combining race categories (white, African-American, Asian/Pacific Islander and Native American) with the ethnic category of Hispanic (implying all race identities are non-Hispanic). The 1990 Census provides a non-Hispanic breakdown of race identical to PGCC race categories, along with an Hispanic/Any Race variable. But for the 2000 Census a race multiple choice format was adopted yielding 62 separate categories! Fortunately, the great majority of Americans tallied by the Bureau opted for a single race identity, and we found recombining categories to correspond with PGCC's format relatively easy. As in the 1990 case, the 2000 race variable was available in a non-Hispanic version permitting the construction of a five-point consolidation of Race/Ethnicity. The Census 2000 SF1 included ethnic Hispanic and non-Hispanic race variable versions restricted to those aged 18 years or older, but the equivalent Census 1990 data was available only in the early release PL 94-171 summary file, not found in the *American FactFinder* data collections. Luckily, this proved to be available from the Maryland State Data Center.

Comparing county residents and PGCC students on income was the most difficult methodological problem we faced. The ideal circumstance would be to have adult *individual* income variables for both county residents and PGCC students. However, the only identical income measure used in both Census years was annual *household* income, and the college does keep accessible student income data in any fashion. Our solution, on both sides, was to measure income indirectly – in the form of a high/high medium/low medium/low scale developed out of Census tract *median* annual household income data. County tracts (both those of 1990 and 2000) were sorted according to their median incomes to identify sensible scale cutting-points (1990 dollars were CPI-adjusted to 2000 values), and then adult residents and PGCC students were distributed across scale intervals according to their home tracts. Student home tracts were determined by address analysis using a geo-coding engine available from *AccuMail*. The pertinent Census income data came from the 1990 SFTP and 2000 SF3 *American FactFinder* collections.

Regionalization of the county also required some set-up data work. To begin with, we needed a scheme of wide applicability, deciding upon a locally well-known division worked out by the Prince George's County Planning Commission. This breaks down the county into seven *planning zones* for purposes of economic development and infrastructure planning. Census and tract-identified student data was added to the planning zones by modeling zones in terms of Census tracts, determining the tract-to-zone correspondence by a map overlay procedure.

Findings

Prince George's Community College's basic market pattern for the decade of the 1990s is easily told (Table 1): While the county experienced moderate growth in its adult population between Censuses (+6.5%), the school was losing enrollment ground, as it turned out, by exactly the same proportion; correspondingly, PGCC draw from its standing market dropped from 3.3% to 2.9%, an overall decline in drawing power of 12.2%. This seemingly simple story is really only the summary outcome of a complex weaving of dozens of smaller tales. When the school's market trends were analyzed geo-demographically (Table 2), a much more detailed and useful picture emerged.

Table 1. Basic PGCC Market Circumstances 1990-2000

Type of Market	Populations			Draw Rates		
	1990	2000	% Change	1990	2000	% Change
County Standing Market (adults)	551,323	586,993	+6.5%	3.29%	2.89	-12.2%
PGCC Customer Base (enrollees)	18,166	16,989	-6.5%	Inap.		

Table 2 reveals the full complexity of the countywide market dynamics affecting the college during the last decade with respect to race, age and income. As previously noted Prince George's County adult population increased by 6.5% from the 1990 census to 2000 census, but this growth was not distributed evenly among all race/ethnic categories. Prince George's County minority population grew significantly between the censuses. The African-American population increased by 34%, the Hispanic population by 83%, and the Asian/Pacific Islander population by 21%. The jump in the Prince George's minority population however, was almost offset by a decline in the number of white residents. From 1990 to 2000, the white population of Prince George's county decreased by 35%.

Similar trends emerged when resident age ranges were considered. The total number of adult residents in the 35-54 and 55 plus age categories actually increased by 22% and 26% respectively, but the lower age categories of 18-24 and 25-34 experienced a 12% and 17% decline. Certainly, the age disparities have implications for future enrollment patterns but also the viability of the county itself. The county is aging and losing a significant number of its younger adult residents. As one would expect, shifts in county race/ethnic and age group demographics affects Prince George's potential market of services needed by individuals. The PGCC unduplicated headcount column shows a 63% decrease in enrollment by white residents and increasing enrollment among its minority citizens. Among age categories enrollment declined for the 18-24 and 25-34 year olds, traditionally the main suppliers of PGCC credit students, while it increased among residents aged 35 years or more.

There was also considerable change in income levels of Prince George's county residents since the 1990 census. The county saw a slight increase (16.4%) in residents making $75,000 or over. Yet, there was a considerable drop in the number of residents making $60,000 to $74,000 thousand (11.9%) and $45,000 to $59,000 (22.0%). The largest gain (69.8%) in resident income appears in the $45,00 or below income group.

Table 2. PGCC Sub-Market Trends 1990-2000

		County Adult Population		***% Chg***	***PGCC Undup Headcount***		***% Chg***
		1990	***2000***	***90>00***	***FY91***	***FY00***	***90>00***
COUNTS	***Total***	*551,323*	*586,993*	*6.5*	*18,166*	*16,989*	*-6.5*
Race/ Ethnicity	White	246,363	160,932	-34.7	7,423	2,720	-63.4
	African American	261,041	349,414	33.9	9,359	12,737	36.1
	Hispanic	21,306	38,971	82.9	401	472	17.7
	Asian/Polynesian	20,518	24,909	21.4	907	941	3.7
	Other/Mixed	2,095	12,767	509.4	76	119	56.6
Age Groups	18-24	94,935	83,884	-11.6	8,669	7,243	-16.5
	25-34	151,156	126,268	-16.5	5,264	4,562	-13.3
	35-54	202,390	247,550	22.3	3,581	4,485	25.2
	55+	102,842	129,291	25.7	651	700	7.4
Income Groups	$75K+	91,422	106,391	16.4	4,455	3,379	-24.2
	$60-74K	163,100	143,621	-11.9	6,969	4,742	-32.0
	$45-59K	182,024	142,044	-22.0	4,756	4,250	-10.6
	<$45K	114,777	194,937	69.8	1,987	4,618	132.4
PERCENTS							
Race/ Ethnicity	White	44.7%	27.4%	-38.6	40.9%	16.0%	-60.8
	African American	47.3%	59.5%	25.7	51.5%	75.0%	45.5
	Hispanic	3.9%	6.6%	71.8	2.2%	2.8%	25.9
	Asian/Polynesian	3.7%	4.2%	14.0	5.0%	5.5%	10.9
	Other/Mixed	0.4%	2.2%	472.4	0.4%	0.7%	67.4
Age Groups	18-24	17.2%	14.3%	-17.0	47.7%	42.6%	-10.7
	25-34	27.4%	21.5%	-21.5	29.0%	26.8%	-7.3
	35-54	36.7%	42.2%	14.9	19.7%	26.4%	33.9
	55+	18.7%	22.0%	18.1	3.6%	4.1%	14.9
	Mean Adult Age	*40.6*	*43.3*	*6.7*	*29.4*	*31.0*	*5.4*
Income Groups	$75K+	16.6%	18.1%	9.3	24.5%	19.9%	-18.9
	$60-74K	29.6%	24.5%	-17.3	38.4%	27.9%	-27.2
	$45-59K	33.0%	24.2%	-26.7	26.2%	25.0%	-4.4
	<$45K	20.8%	33.2%	59.5	10.9%	27.2%	148.5
	Med HH Income	*$56,841*	*$55,256*	*-2.8*	*$61,362*	*$58,257*	*-5.1*
DRAW RATES							
Race/ Ethnicity	White	3.01%	1.69%	-43.9			
	African American	3.59%	3.65%	1.7			
	Hispanic	1.88%	1.21%	-35.6			
	Asian/Polynesian	4.42%	3.78%	-14.5			
	Other/Mixed	3.63%	0.93%	-74.3			
Age Groups	18-24	9.13%	8.63%	-5.4			
	25-34	3.48%	3.61%	3.7			
	35-54	1.77%	1.81%	2.4			
	55+	0.63%	0.54%	-14.6			
Income Groups	$75K+	4.87%	3.18%	-34.8			
	$60-74K	4.27%	3.30%	-22.7			
	$45-59K	2.61%	2.99%	14.5			
	<$45K	1.73%	2.37%	36.8			

The lack of parity in income, age and race/ethnicity groupings can affect market draw rates. In 1990, PGCC enrolled a little over 3% of the county white residents. By the 2000 census the rate fell to 1.69%, which represents a 44% enrollment loss. While the African-American population actually increased, the college's draw rate for this ethnic group remained constant. The market draw rate for Hispanic students actually decreased by 36%, despite an increase in the Hispanic adult population. A similar result emerges for Asian/Polynesian and mixed other race/ethnicity categories. The Hispanic, Asian/Polynesian and mixed other populations grew tremendously by the 2000 census, but the college attracted fewer of these students to its programs.

An examination of draw rates by age and income groupings reveal comparable trends. Five percent fewer residents aged 18-24 enrolled at PGCC in 2000 than 1990. This is not a surprise given the overall population decline in this age group. Despite a decreasing number of adults in the 25-34 age category, PGCC managed to maintain a constant draw rate for this group. In the 35-54 age group, the college increased its draw rate by around 2%. Among residents 55 years and older residents the college's market rate declined by 15%, suggesting a need for more senior level activities.

The greatest variability in draw rates occurred among the varying income groups. The college draw rate for adult residents earning $75,000 or more declined by 35%. Likewise, the draw rates for residents earning $65,000 to $74,000 declined by 23%. There could be a myriad of factors that pushed the market draw rate downward among these income groups. But, in general individuals at higher income levels have more education and do not need require additional educational services. The largest increase in PGCC draw rates by income occurred among residents earning less than $45,000 or less. The college's draw rate among this group has kept pace with the population growth for this income bracket.

Table 3 shows what happened when we extended our market analysis to county geography. It displays the PGCC basic market trends for census years 1990 and 2000 by each of the seven individual county planning zones and their groupings into two global categories: residents inside and outside the Washington Beltway. The inside-the-Beltway zones are considered the more urban centers of the county. The outside-the-Beltway zones are considered the more suburban regions of the county. PGCC uses planning zones data on demographics and market draw rates in identifying underserved areas and populations and for planning ameliorating new extension centers and instructional programs.

As can be seen, there was only a slight decrease (1.4%) in the number of county adults living inside the beltway, traditionally our best source of enrollment, from census to census. The same period also saw a slight decrease (1.0%) in the number of enrolled PGCC residents from this super-region, comprised of planning zones 2, 4, and 7. The Inner Beltway draw rate remained static during this interval, although it deteriorated significantly in one of its constituents planning – Planning Zone 7 – by over 13%. In contrast, the Outer Beltway market (zones 1, 3, 5 and 6) grew by over 20%, even while PGCC's draw from this now much richer potential source of students dropped by a worrisome 27%. Currently, the two most underserved planning zones turned out to be Zones 1 and 2 on both sides of the Beltway in the county's northern reaches.

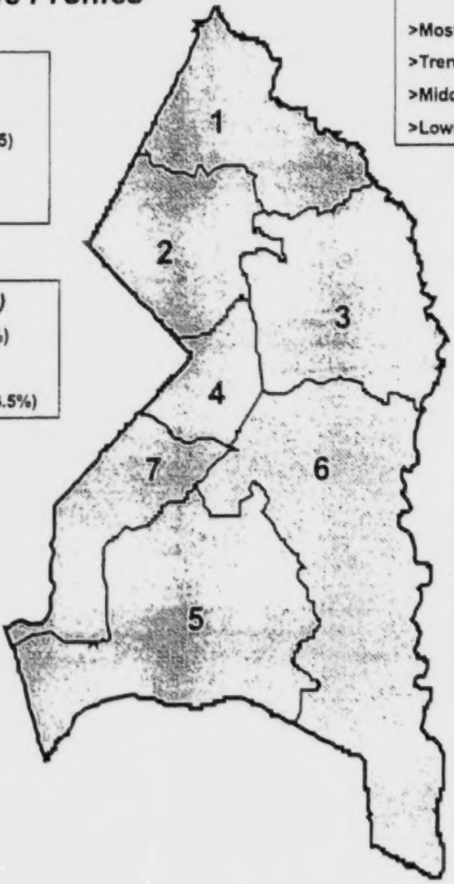
Table 3. PGCC Market Trends 1990-2000 by County Planning Zones

Planning Zones	*County Adults*			*PGCC Students*			*Draw Rates*		
	1990	**2000**	**% Chg**	**FY91**	**FY00**	**% Chg**	**FY91**	**FY00**	**% Chg**
Inside Beltway	*351,313*	346,404	*-1.4*	*9,529*	*9,437*	*-1.0*	*2.7%*	*2.7%*	*0.4*
2 (Northwest)	189,744	185,227	-2.4	4,133	4,195	1.5	2.2%	2.3%	4.0
4 (W Central)	78,064	79,557	1.9	2,480	2,769	11.7	3.2%	3.5%	9.6
7 (Southwest)	83,505	81,620	-2.3	2,916	2,473	-15.2	3.5%	3.0%	-13.2
Outside Beltway	*200,010*	240,509	*20.2*	*8,637*	*7,552*	*-12.6*	*4.3%*	*3.1%*	*-27.3*
1 (Northeast)	58,467	65,685	12.3	1,584	1,241	-21.7	2.7%	1.9%	-30.3
3 (E Central)	69,454	86,766	24.9	3,642	3,133	-14.0	5.2%	3.6%	-31.1
6 (Southeast)	28,322	35,468	25.2	1,432	1,366	-4.6	5.1%	3.9%	-23.8
5 (S Central)	43,767	52,590	20.2	1,979	1,812	-8.4	4.5%	3.4%	-23.8
Total	*551,323*	*586,913*	*6.5*	*18,166*	*16,989*	*-6.5*	*3.3%*	*2.9%*	*-12.2*

Limited space does not permit a full display or treatment of planning zone market demographics, but the capsule profiles of each zone's demographics and market behavior provided in Figure 1 should prove a sufficient basis for a sketch of their main qualities. All three Inner Beltway zones have high concentrations of African American residents (over 65%), younger adult populations compared with Outer Beltway zones, and residents who tend to live in neighborhoods characterized by sub-affluent households. Zone 2 stands out for also being the home of the highest proportion of Hispanic and Asian adults (22%), the highest proportion of young adults (21%) and for being the second lowest proportional geographic source of PGCC enrollment (2.3% draw).

The four planning zones outside the Beltway, which used to be almost exclusively white racial preserves, now manifest something akin to racial balance, the result of substantial white outmigration over the last decade. While no longer as affluent as they were in 1990, they continue to be more prosperous than the Inner Beltway zones and to house more aging populations. Three of the four (3, 5, and 6) register above average PGCC draw rates, although, as has already been mentioned, draw rate on the whole has declined sharply outside the Beltway since 1990. Zone 5 (South Central) is currently the wealthiest and most aged region (93% of its adults live in tracts with median household incomes exceeding $75,000 and 60% are over 40 years old). Zone 6 (Southeast) is special for its very high PGCC draw rate (3.9%), due in part, to be sure, to the fact that the college is located here. But the Outer Beltway area most interesting to PGCC planners is Zone 1 (Northeast). It represents the highest geographic concentration of white and Asian adults in the county (49% and 8%, respectively), but manifests the least inclination to send students to the college (1.9%).

Figure 1. Planning Zone Profiles



Northwest (-2.4%)

>Most Hispanic/Asian (22%)

>Youngest population (21% Age <25)

>Majority low income (54%)

>2nd lowest Draw Rate (2.3%)

West Central (+1.9%)

>Most African American (90%)

>Majority low income (55%)

>Above average Draw Rate (3.5%)

Southwest (-2.3%)

>Largely African American (80%)

>2nd oldest zone (57% Age 40+)

>Most low income (77%)

>Moderate Draw Rate (3.0%)

South Central (+20.2)

>Fair white/black ratio (29%/63%)

>Most aged area (60% over 40)

>Almost all high income (93%)

>Above average Draw (3.4%)

Northeast (+12.3%)

>Most white (49%) and Asian (8%)

>Trending middle aged

>Middle/High income mix

>Lowest Draw Rate (1.9%)

East Central (+24.9%)

>Good white/black ratio (37%/56%)

>Trending middle aged

>2nd in high income (90%)

> 2nd highest Draw Rate (3.6%)

Southeast (+25.2%)

>Good white/black ratio (37%/56%)

>Trending middle aged

>73% high income

>Highest Draw Rate (3.9%)

In the past two years PGCC has opened two new extension sites based on planning zone data trends. Our first site, the Metro Center opened in the Fall of 2000 in Zone 2, and we are happy to report that healthy enrollment growth has occurred every term since inception. The college's second new site, the Laurel Center, opened in the spring of 2002 in a Zone 1 location. Although the period since its establishment is too recent for proper assessment, preliminary data on enrollment suggest considerable potential.

Table 4. Result of Establishing PGCC's New Metro Center
(Unduplicated Combined FY01 and FY02)

	% County Adults	% Non-Metro Center	% Metro Center Students	Ratio: Non-Metro /County	Ratio: Metro /County	Ratio: Metro/ Non-Metro
All Zones						
Inside Beltway	*59.0%*	*54.6%*	*78.4%*	*.92*	*1.33*	*1.44*
2 (Northwest)	31.6%	22.9%	68.2%	.72	2.16	2.98
4 (W Central)	13.6%	16.9%	5.9%	1.24	.43	.35
7 (Southwest)	13.9%	14.8%	4.3%	1.06	.31	.29
Inner Beltway	*41.0%*	*45.4%*	*21.6%*	*1.11*	*.53*	*.47*
1 (Northeast)	11.2%	7.5%	10.3%	.67	.92	1.37
3 (E Central)	14.8%	19.8%	7.7%	1.34	.52	.39
6 (Southeast)	6.0%	8.5%	1.5%	1.42	.25	.18
5 (S Central)	9.0%	9.8%	2.1%	1.09	.23	.21
(N)	(589,913)	(24,246)	(1,912)			
Zone 2 Only						
White	28.3%	15.1%	8.8%	.53	.31	.58
Black	47.5%	76.8%	73.7%	1.61	1.55	.96
Hispanic	15.7%	2.8%	10.9%	.18	.69	3.87
Asian	5.8%	4.6%	5.9%	.79	1.02	1.28
(N)	(185,227)	(1,582)	(491)			

Table 4 provides data bearing on the success of the college's Metro Center venture. The first column identifies Center enrollment (any Center-located course taken) across all planning zones. The second set of results reports the race/ethnic enrollment percentages of students who exclusively reside in Zone 2 where the metro center is located. The data indicate that the great majority of students at the Metro Center reside inside the beltway, with 68 percent coming from planning zone 2 alone. Only 22 percent of Metro Center enrollees reside outside the beltway in planning zones (1, 3, 6 and 5). These outcomes are consistent with the college's goal of drawing more students from Zone 2 by establishing a viable instructional facility within its borders of this most populous region.

Furthermore, the race/ethnicity distributions shown in Table 4 are also hopeful. The secondary goal for establishing an extension center in Zone 2, a haven for recent immigrants in the county, was to attract more Hispanic and Asian students into the college's classrooms. Although the single largest racial group represented in Metro Center's two year enrollment was the African American (74%), 11% turned out to be of Latino background, this proportion was almost four times (3.87) the size of the Hispanic component of the non-Metro Center student body for this time interval. Likewise, although less dramatically, Asians enrolled at the Center at a rate 28% greater than they did at other college instructional facilities. These findings fail to prove the effectiveness of the college's recruitment strategy (among other things, apparent enrollment lifts at a new extension center may represent a re-distribution of current students rather than the attraction of new students), but at least the data patterns provide some ground for optimism.

We conclude this paper with the expression of a hope and an intent. The hope is that the methodological and substantive findings presented here will encourage other institutional researchers to approach enrollment analysis in a fashion that integrates student record data with service area demographic data. The intention is our plan to deepen our analysis in future market studies to include marketing dimensions left unexplored in this preliminary research. The short list includes coming to grips with the role of the county's practical and ready markets of higher education in conditioning PGCC enrollment, and to bring market share data into the analysis in order that draw rate gains and losses can be assessed as possible products of the recruitment successes and failures of rival postsecondary institutions with which we share our service area. We are well aware that we have much more work to do before we achieve our goal of achieving a comprehensive understanding of Prince George's Community College's marketing environment and its impact on the size and composition of its student body

References

Baldwin, A. (1998). The Draw to Florida Community Colleges from Florida's "College-Age" Population by Program Areas. Office of Institutional Research, Miami-Dade Community College. Miami, FL.

Boughan, K., Diehl, P. (1995). Lifestyles of the Targeted and Enrolled: Using Geo-Demographic Analysis at the Community College. Paper presented at the annual forum of the North East Association for Institutional Research, Burlington VT.

Heacock, R., Jenkins, J., (1993). Howard Community College Enrollment by Census Tract, Fall 1991. Research Report Number 81. Columbia, MD.

Hecht, A. (1991). Student Ethnic Diversity at Parkland College. Champaign, IL [ED335079]

National Center for Higher Education Management Systems (1998). Analysis of Potential Needs for Postsecondary Education Services in Different Regions of Oklahoma. Boulder, CO.

Newmyer, J., Mangham, C. (1992). Proposals for: New College for the Los Rios Community College District and New Center for the Allan Hancock Joint Community College District. Los Rios, CA. [ED351072]

American FactFinder (2002), U.S. Census Bureau Internet site: Factfinder.census.gov/

Virginia State Department of Community Colleges (2000). Population Participation Rates by City and County, Fall Term 2000. Virginia Community College System Research Report Series. Richmond, VA, 2000.

LINKING ASSESSMENT PLANNING AND MISSION REVIEW: ONE UNIVERSITY'S EXPERIENCE

Ellen Boylan-Fick, M.P.S.
Assistant Director of Institutional Research
Barbara R. Sadowski, Ph.D.
Assistant to the President for Planning and Research

Marywood University
Office of Planning and Institutional Research

Introduction

Ideally, the process of developing an assessment plan begins with rigorous examination of an institution's mission statement. The conceptual elements set forth in the mission statement lay the foundation for institutional goals and objectives that shape all the work of the university, from the classroom to the weight room, the boardroom to the lab. Most importantly, though, the mission statement lays the foundation for the institutional assessment plan and sets the criteria against which it is tested.

The task of assessment planning, whether it is done in the Institutional Research office or not, is to build a culture of assessment that is evident in every part of the university community. There are two good reasons for this: to prompt change and improvement where needed, and to serve beautifully as concrete evidence of successful adherence to assessment standards set by accrediting agencies.

Objective

The primary objective of this paper is to describe the role of one institutional research office in a venture to build an institutional assessment plan and revise the mission statement at the same time, activities so closely linked that one cannot proceed absent the other. This will help institutional researchers envision their own role and responsibility in supporting and promoting ongoing assessment activity throughout the university, and help them ensure that a link back to the institution's mission, goals, and objectives is preserved.

Building a Culture of Assessment

History of Assessment Planning

Outcomes assessment at Marywood University, a small private university in northeastern Pennsylvania, became a campus-wide endeavor as the 1996 Middle States Accreditation (MSA) visit came ever closer. At that time, the MSA standards called for a campus-level coordinating body with oversight responsibility for learning outcomes assessment. The membership of the initial Outcomes Assessment Group (OAG) consisted of the deans of the four schools, a faculty member from each school, a representative from each of the four vice-presidential areas, the enrollment management administrator, and the assistant to the president for planning and research as chair.

The OAG met for the first time in March 1996, without a clear mandate other than to document outcomes assessment activities currently in place on campus. Although the OAG met once a semester thereafter, little progress was made except for a modest list of learning outcomes measures from departments and programs which, for the most part, had external accrediting visits that required documentation of learning outcomes. Examples of external criteria, or direct evidence, for measuring learning outcomes included the results of licensure examinations for education, nursing, dietetics, and the like. Other departments such as Social Work, Art, Speech Pathology, Music and Business required written evidence of learning outcomes assessment as part of reports to the accrediting body.

The IR office collected and organized the data about learning outcomes assessment, summarized it and reported back to the OAG each year. In addition, indirect evidence of student satisfaction based on data from surveys conducted by IR such as the annual senior survey and the campus diversity survey were shared with the OAG. Reports from other areas of the university such as career services' employment report, university advancement's capital campaign updates and enrollment management reports were listed as evidence of institutional effectiveness.

In 1998-99, the OAG assisted in the selection of the quadrennial student satisfaction survey administered in the fall of 1999. The results of that satisfaction survey were presented to several campus groups following data analysis. Over the next year, some IR staff members attended workshops on assessment at professional meetings such as NEAIR and SCUP, notably pre-conference workshops by Fred Volkwein of The Pennsylvania State University. Although faculty members on the committee endorsed the idea of bringing in an assessment consultant to provide professional development for faculty members, a time and resources could not be identified to make it happen.

In 2000-01, as the MSA Periodic Review Report (PRR) loomed on the horizon, the focus on outcomes assessment intensified. A written report detailing current outcomes assessment was submitted as part of the PRR, yet there was little buy-in from the faculty as a group or departments that had no external body requiring outcomes assessment. Very modest progress had been made toward creating a culture of assessment on campus despite the five-year history of the OAG. All this was about to change.

Impetus for Change

The most important factor in changing the climate of outcomes assessment at Marywood University was leadership and support from the highest administrative level. It began in the summer of 2001. Prior to the start of the 2001-02 academic year, Marywood University's president mandated that a draft of an institutional outcomes assessment plan be written and submitted to her by June 2002. One of the first steps was to make the Outcomes Assessment Group a university standing committee, the Outcomes Assessment Committee (OAC), reporting directly to her and coordinated by the staff of planning and institutional research.

The responsibility for writing the policy and procedures for the OAC fell to the head of planning and institutional research. As a university standing committee, the purpose, membership, and reporting lines of the Outcomes Assessment Committee had to be submitted to the Committee on Committees and then to the university policy committee for approval. The purpose of the OAC was to advise the president on issues related to outcomes assessment including recommendations about resources needed to support the process, and to coordinate the process of developing the campus outcomes assessment plan.

Unlike some institutions, Marywood University's definition of outcomes assessment included more than just student learning outcomes; it also included evidence of institutional effectiveness. As a result, the OAC membership included representatives from all areas of the university, including business affairs, university advancement, student life, academic support areas and all academic units. All academic deans are committee members as is the Associate Vice President of Academic Affairs - Enrollment Management. The Director of Human Resources represented business affairs, the Assistant Vice President for Development represented University Advancement, and the Director of Career Services represented Student Life. Three faculty members were appointed by their respective deans. All vice presidents were *ex officio* members of the committee. Figure 1 shows the relation between the Office of Planning and Institutional Research and the Outcomes Assessment Committee. A comment on the role of Institutional Research in assessment is necessary to fully understand how the process worked at Marywood.

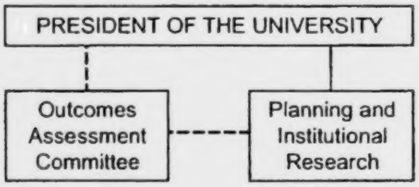

PRESIDENT OF THE UNIVERSITY
Outcomes Assessment Committee
Planning and Institutional Research

Figure 1

Role of IR in Assessment

At Marywood University, planning and institutional research functions are integrated as a single department (P&IR) reporting to the president. As such, the P&IR office tracks all institutional data and has an institutional focus that is also linked to strategic planning. Not only does this office coordinate all institution-wide surveys and prepare reports, it tracks benchmarks on progress towards the strategic goals. The IR staff also responds to data requests from other senior administrators, prepares reports on faculty productivity, maintains a password-protected EIS (Executive Information System) for different levels of administrators, and a website on the campus intranet for disseminating information to faculty and staff.

Because of the position of planning and institutional research within the institution and staff expertise in measurement, it seemed logical to give the IR office the responsibility for coordination of outcomes assessment. Not every IR office reports to a president, nor does an IR staff necessarily have the requisite background to lead others in developing an assessment plan. IR professionals have more experience in assessment than most other campus units. By recognizing that the surveys managed by IR often produce indirect self-reported evidence of student learning outcomes, IR serves an important role in contributing to an overall institutional assessment plan.

Year One Activities

The president's mandate to produce a written draft of an assessment plan by June 2002 was accompanied by financial resources to assist professional development of OAC members. Funds were provided by the president. One dean and the head of P&IR attended a pre-conference workshop and Assessment Conference held in Indianapolis in November 2001. IR staff attended workshops (Volkwein, 2001), as well. The conference attendees gathered a wealth of information about assessment including vendor information about commercial resources such as TrakData, possible consultants to assist faculty in developing assessment plans, assessment newsletters, websites and many examples of assessment models, instruments and plans. Many of the conference attendees and presenters were senior faculty members who hold leadership roles in assessment planning. At Marywood, however, no clear faculty leader had been identified at that time.

Attendance at two MSA assessment conferences during the fall of 2001 were also funded to help committee members understand what was required as part of an outcomes assessment plan. The fall undergraduate faculty forum at Marywood was devoted to a presentation on the new MSA standards on outcomes assessment. Meetings were held in each vice-presidential area to discuss assessment plans. The chair of the OAC was invited to several of these area meetings to help departments define exactly what assessment meant for the area. In some instances, the Continuous Quality Improvement (CQI) process was identified as the appropriate model for thinking about their unit's role in assessing institutional effectiveness. By late fall, the deans had decided that they would be the assessment leaders for their schools, rather than appointing a faculty member as the campus leader for assessment of learning outcomes. Each dean had the group of department chairs work on assessment planning for programs and departments.

While the year-end goal was to have a draft of an outcomes assessment plan, the process was divided into four stages that would take place over two years, in order to avoid unnecessary rewriting. The first stage, introduced in that fall of 2001, would be to complete the inventory of existing OA activities in each unit and place these in context by introducing a model for an assessment plan. The second stage would add information about the timetable for implementation, products and dissemination of assessment results, completing another part of the model. In the third stage, at the start of fall 2002, goals and objectives would be written for each unit. In the fourth stage, evidence of actions and feedback from the assessment process would be added to the plan. A brief digression to explain why the staging was necessary for Marywood University follows below.

A Concurrent Restructuring of Academic Affairs

Marywood University was in the process of restructuring academic affairs from schools into colleges as the assessment planning process began in earnest in 2001. A year later, decisions had been made about which departments and programs would be aligned in the four new colleges. The plan was that by June 2003, restructuring would be formalized.

The 2002-03 academic year was a transition year between the old and new structures, and unit goals and objectives linked to institutional goals and objectives were not yet set down. This held up the assessment planning process, because asking faculty and administrators of the former schools to develop new school goals and objectives did not make sense. So, the logical first step of deciding on goals and objectives for the assessment plan was delayed until the faculty was working together on the missions, policies and procedures for the new colleges during the 2002-03 transition year.

A second important reason for the delay was the on-going mission review process. The timing of the request to write unit goals and objectives was critical because a campus-wide committee chaired by the president was in the process of reviewing the mission and derived institutional goals and objectives. Without a clear consensus about institutional goals and objectives, school/college goals and objectives would be difficult to develop without rewriting. The work of the mission review committee is described below.

The Mission Review Process

As the form and function of outcomes assessment on campus was gaining definition during that first year of committee action, our mission statement was set to undergo change, as well. It had become clear to the President, who sits on the MSA Board, that our stated mission needed revision in order to better reflect the university we had become during more than a decade of stunning advancement and renewal since 1990. She established the Mission Review Committee in Fall 2001 to undertake that task.

The previous year the campus had participated in a SWOT analysis as part of strategic planning, and also had developed a consensus about the core values of the university. Committee work began by considering the adequacy of the current mission to meet the criteria for a mission set out by new MSA standards. Several articles on what a mission should be, as well as examples of mission statements from similar institutions were examined. Lively discussions centered on major changes at Marywood since the last mission review, and whether the current mission adequately reflected the changed institution. For example, since the last mission review, Marywood had become fully coeducational, strengthened sports and athletics, attained university status and inaugurated two doctoral programs.

Membership on the Mission Review Committee included faculty from a variety of disciplines, the undergraduate and graduate deans, and administrators from university advancement, planning and research, and business affairs. A strategy for conducting the mission review was loosely based on a model described on the Meeting Facilitators International web page (Withrow, 2002). Committee members then determined that a strong mission statement would develop from five *assumptions*, and needed to address five questions, or *query items*:

A Model for Mission Review

Assumptions	***Query Items***
Answers the query items	*Who we are*
Addresses our primary audiences	*Whom we serve*
Connects and reflects core values	*What we do for them*
Is succinct	*How we do it*
Avoids redundancy	*Our uniqueness*

Figure 2

The small but earnest group of committee members set down some initial responses to the queries, but it soon became apparent that broad-based input would inform the document most eloquently. One characteristic about Marywood University that is unique is a history of collegial engagement in issues important to the operation of the institution, and our president continues to be true to that legacy. As a consequence, staff of the Office of Planning and Institutional Research was called on to conduct a *qualitative study* consisting of focus group research on mission, using a cross-section of campus representatives.

The interviews were conducted over a month-long period in the spring semester. Members of the Mission Review Committee were recruited to lead individual focus groups. Instructions and work sheets for leaders and group members were created in order to help the participants engage more fully in the discussion and ensure that responses were elicited, at least loosely, by query categories. Focus group discussions were taped and later transcribed by institutional research staff.

Text of the transcriptions from each focus group was combed for raw responses, listed and later grouped by query item. Common phrases and themes that emerged from the text were set down and then submitted to the Mission Review Committee for discussion. These phrases and themes will augment and enrich the expression of mission that the committee eventually crafts and shares with the university community.

Goals and Objectives Developed

Late in spring 2002, the current institutional goals were revised and rewritten to better reflect the revised mission. Feedback from campus leaders resulted in a second revision to the draft of goals and objectives in early fall. By mid-fall the draft of the core values, proposed mission, goals and objectives was disseminated to all areas of the campus for review, discussion and comment. A writing committee was formed to review and consolidate all comments from discussion groups.

As this was going on, early steps to develop the assessment plan, such as creating an inventory and finding models to guide our work, were proceeding simultaneously. The relation between the mission review process and assessment planning stages are shown in the schematic in Figure 5.

Developing the Assessment Plan

As noted above, completing the assessment inventory was the first task assigned to each campus unit. Background articles like the one on MSA standards by George

Santiago, Jr. (2001) in *Assessment Update* were distributed to OAC members, and models like one from The University of St. Thomas (1999) were introduced to help give structure to the task of developing an assessment plan. One very helpful assessment model was provided by Trudy Banta, Vice Chancellor for Planning and Institutional Improvement at Indiana University – Purdue University Indianapolis, and Editor of *Assessment Update.* Her "Planning for Learning and Assessment" questions are shown in Figure 2.

PLANNING FOR LEARNING AND ASSESSMENT
T. Banta (2002)

1. What general outcome are you seeking?	2. How would you know it? (the outcome) if you saw it? (What will the student know or be able to do?)	3. How will you help students learn it? (in class or out of class)	4. How could you measure each of the desired behaviors listed in #2?	5. What are the assessment findings?	6. What improvements might be based on assessment findings?

Figure 2

Using this model as a foundation, we developed our own, one that would be useful for making a general inventory of activities for institutional assessment overall, not just student learning outcomes. Our Marywood "Assessment 5-Year Report" is shown in Figure 3.

ASSESSMENT 5-YEAR REPORT FOR 0000 - 0000
(Marywood University, 2002)

1. Intended Outcome or Objective of the Major or Program	2. Assessment Method	3. To Whom /When	4. Product, and Date	5. Oversight Body	6. Actions Taken on Results

Figure 3

Units or departments with an existing assessment plan did not have to rewrite to conform to the model as long as all relevant information was included. The office of institutional research provided an example of information to include as part of current assessment activities. Stage one of unit plans were due in March 2002, when each area would report back to the committee.

The second stage involved completing information about what was produced as a result of the assessment and to whom these results were disseminated. For example, the senior survey results are distributed in a report produced every other fall with data posted to the Marywood intranet, a complete report to the cabinet and summary presentations to the faculty and other administrative units. At the spring OAC meeting, information about the new MSA standards was distributed, along with pertinent articles from *Assessment Update*. A template of a budget summary spreadsheet was distributed to help track unit resources devoted to assessment. Planning began for a consultant visit for academic affairs in the fall. Unit assessment plans with stage one and two completed were due by June 1, 2002. In reality, assessment plans were all turned in by the end of the summer.

The third stage (planned for fall 2002) of developing the assessment plans required each unit to write unit goals and objectives that were *linked to institutional goals and objectives*. Note that institutional goals and objectives are not the same as strategic goals and objectives.

Strategic goals and objectives derive from institutional goals and objectives. Because strategic goals may change from one planning cycle to the next, they can be viewed as aspects of the institutional goals that receive greater attention or emphasis for a short period of time in response to external events and conditions. MSA standards require a clear link between the mission statement and institutional goals and objectives as well as links between the institutional and unit goals and objectives.

Year Two Activities

Now, in year two of this venture at Marywood, an assessment consultant has already presented to faculty members and department chairs. His day-long consultancy also involved working with student life administrators answering questions about their role in contributing to student learning outcomes. Staff in business affairs attended an EACUBO meeting and identified an expert in assessment who described how she had managed the assessment process for fiscal affairs. She will serve as a consultant to the departments in business affairs as the year progresses. University advancement and student life have also been offered the services of a consultant in their respective areas to help with assessment.

A draft of the *institutional* goals and objectives was distributed to members of the OAC at the fall meeting. The primary goal for assessment planning during the second year is to write goals and objectives for each department or unit and then determine if current assessment activities are adequate. Once the preliminary unit goals and objectives are written, each unit will then evaluate their current assessment activities in light of them. Information based on assessment activities should enable the department or unit to determine if goals have been achieved or if changes are necessary.

A culture of assessment results in a continuous process of setting goals, measuring whether progress has been made or a goal achieved, making changes and then assessing whether the change has been effective in moving the unit closer to achieving its goals. The data from an assessment should provide information to help support decisions about changes needed to progress towards a goal. So, the last stage of developing assessment plans is to identify whether additional assessment measures are necessary, what the cost is for the assessments and when they will be administered. For some departments, evidence of improvements resulting from prior assessment activities will complete the plan

As an example, an institution may have a goal of providing a challenging educational environment for its students. One possible method to assess progress towards the goal is the National Study of Student Engagement (NSSE) benchmark called Level of Academic Challenge. The data from the NSSE is compared to both what might be expected of students at a given institution as well as compared to other institutions nationally from the same group. The Level of Academic Challenge benchmark is made up of several items asked of both seniors and first year students about class experiences such as the number and length of papers, class presentations, being asked to work harder than they thought they were capable of doing, etc. By examining the student responses to the individual items, feedback about a student's total educational experience can provide faculty with directions for curricular changes. A second administration of the NSSE can then provide evidence of whether the changes have had the desired effect.

An Ongoing Venture

In recent years, efforts to establish ongoing assessment have become more ardent in academe, and standards written by the Middle States Association have called ever more forcefully for a link between assessment and mission, goals and objectives. Marywood University is now engaged in a two-year venture to respond to these pressures, and an assessment plan based on a strong mission statement will be soundly in place by the close of this academic year. The finished plan, drawn from all quarters of the university, and involving faculty members, staff and administrators, will be solid evidence of the link between assessment and mission.

A culture of assessment is being bred at Marywood University. It will support an ongoing cycle of setting goals, measuring progress, making change, and evaluating the effectiveness of that change in achieving institutional goals. The most important role we play in institutional research is to help establish a process of assessment that promotes "overall institutional improvement" (MSA 2002), a process that continues for the life of the university.

Building a Culture of Assessment

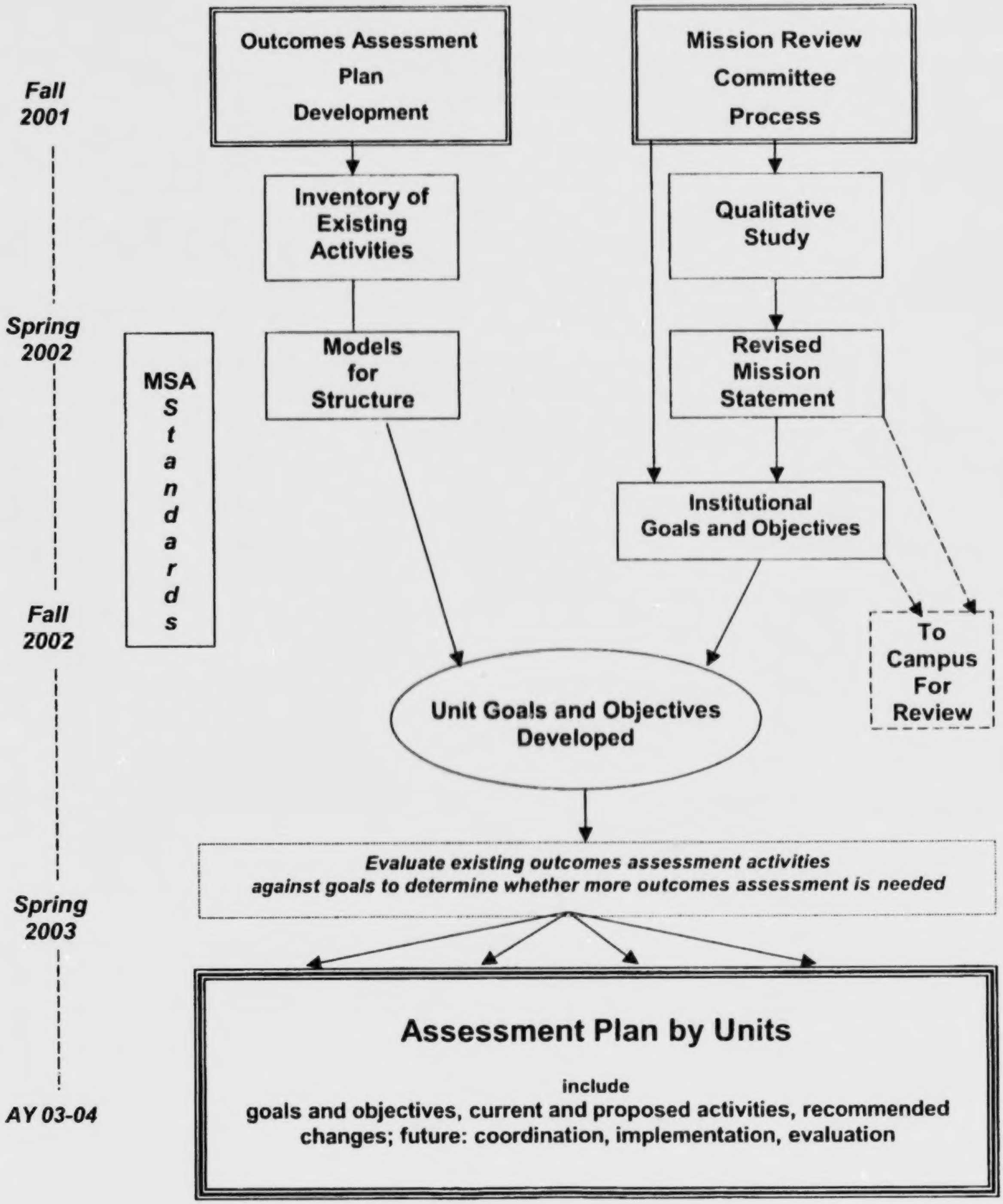

Fall 2001
Outcomes Assessment Plan Development
Mission Review Committee Process
Inventory of Existing Activities
Qualitative Study
Spring 2002
MSA Standards
Models for Structure
Revised Mission Statement
Institutional Goals and Objectives
Fall 2002
To Campus For Review
Unit Goals and Objectives Developed
Evaluate existing outcomes assessment activities against goals to determine whether more outcomes assessment is needed
Spring 2003
Assessment Plan by Units
include
goals and objectives, current and proposed activities, recommended changes; future: coordination, implementation, evaluation
AY 03-04

Figure 5

References

Banta, T.W., Ed. *Assessment Update: Progress, Trends, and Practices in Higher Education.* San Francisco, CA: published bi-monthly by Wiley Subscription Services, Inc., A Wiley Company, at Jossey-Bass.

Middle States Association, Commission on Higher Education. (2002). *Assessment of Student Learning: Options and Resources.* Philadelphia, PA.

Santiago, George, Jr. (2001, September-October). Toward Excellence in Outcomes Assessment: The Middle States Approach. *Assessment Update, 13(5).*

The University of St. Thomas. (1999). Practical Tools for Assessment, from a faculty workshop sponsored by the Academic Assessment Coordinating Committee. St. Paul, MN: The University of St. Thomas.

Volkwein, J.F. (2001, November). *Responding to Accreditation and Assessment on Your Campus: Why, What, Who, How?* Workshop presented at the 28th Annual Conference of the North East Association for Institutional Research, Boston, MA.

Withrow, Bruce. (2002). "Developing a Mission Statement." *Meeting Facilitators International.* Retrieved November 12, 2002, http://www.facilitators.com/workshops

~

Assessment Web Sites

http://lsv.uky.edu/archives/assess.html
The oral history of the ASSESS listserv and instructions for subscribing are here.
http://www.departments.bucknell.edu/inst_research/
A library of assessment resources available on web and summary of "best practices."
http://www.chems.com National Center for Higher Education Management Systems (NCHEMS).
http://www.ncsu.edu/undergrad_affairs/assessment/files/evaluation/outline.htm
North Carolina State University: undergraduate assessment plans, models; also, information on their Undergraduate Assessment Symposium April 7-8, 2003 in Raleigh.
http://www.uncwil.edu/stulife/
University of North Carolina at Wilmington has this site for student life assessment.
http://web.umr.edu/~assess/other/instass.html
A list compiled by the University of Missouri, Rolla of assessment pages at more than 15 universities. Missouri's site also features their own model for Co-curricular Assessment of Skills and Education (C.A.S.E.)
http://cortland.edu/oir/assmtpage.html
SUNY Cortland assessment activities
http://pages.towson.edu/assessment/towson_assessment_activities.htm
Assessment reports (plans plus results) for Towson's Chemistry, Environmental Science & Studies, and Theatre programs.

AN ASSESSMENT OF STANDARDIZED ACCUPLACER PLACEMENT SCORES FOR COLLEGE ENGLISHE IN THE CONNECTICUT COMMUNITY-TECHNICAL COLLEGE PROGRAM

Corby A. Coperthwaite, Ph.D.
Director of Planning, Research and Assessment
Board of Trustees
Connecticut Community-Technical College System

William F. Ritchie, Ph.D.
Director of Institutional Research
Tunxis Community College

Abstract

The 12 Connecticut community colleges use Accuplacer for placement into college English. However, each college uses different placement cut-off scores, and there are varying structures of developmental education in place at each school. Recently, a recommendation was made to implement a common placement standard across the system.

This study examines what the impact of the proposed standard would have been on new and transfer students enrolled in English Composition and Developmental English during the Fall 2001 term. The proposed cut scores would not have impacted the cohort of Developmental English students. However, the proposed cut scores would have impacted the placement of English Composition students considerably and suggests, based on limited data, that the proposed cut scores be reconsidered.

Introduction

Basic skills assessment for entering students has become more than an academic policy issue in the Connecticut Community College System. Within the system, each of the 12 colleges uses the Accuplacer for placement into college English; however, each college uses different placement scores, and there are varying structures of developmental education in place at each school. Currently, a student can place into Developmental English at one college in the system and then take those same placement test scores to another Connecticut community college and place into college English. The problem is further complicated when transfer articulation agreements are considered, because college English completers may or may not possess the same proficiency level in the subject.

Recently a recommendation was made to implement a placement standard for college English. The proposed placement scores for college English would require a score of 8 or higher on the Accuplacer Essay, an 83 or higher on Accuplacer Reading Comprehension, and an 86 or higher on Accuplacer Sentence Skills. The purpose of this study is to assess the impact of this recommendation. This study does not address pure learning outcomes; rather it is descriptive and limited to the following research questions: (1) How would the proposed placement scores affect students who placed in college English or Developmental English under the current local college norms? (2) Do the proposed placement scores result in an increase or decrease in the number of developmental students? (3) Do the proposed placement scores result in an increase or decrease in the number of developmental sections? (4) What are the implications, if any?

Methods

Course File

Course data come from the Fall 2001 course table that is part of the third week suite of frozen extract files (SWRXF22) in Banner. Before the file could be used, the following file maintenance was performed. There are a variety of Developmental English courses throughout the community college system. Some are uniquely writing or reading courses while others combine the two skills within a single course or series of courses. Some colleges consider reading as English and some do not. For the purposes of this study, Developmental English will include mainstream writing and reading courses. The study excludes English as a Second Language, English for Deaf Studies, Introduction to Thinking, and others. The first college level English course was defined by the colleges and will be referred to as English Composition throughout this paper.

Student file

Student data come from the Fall 2001 student information table from the SWRXF22.

Testing File

Accuplacer scores come from the Fall 2001 placement test table from the SWRXF22. Because the file records one record per placement test session, a student can be in the file numerous times so a new table was created with one record per student that included the most recent Accuplacer scores. Because the study is predicated on the use of Accuplacer, scores from locally developed and other commercial placement instruments used by some colleges for placement were not considered.

Sample

To be included in the sample a student needed to be a new or transfer student and be enrolled in developmental reading and/or writing or English Composition during the Fall 2001 semester. In addition, the student also needed an Accuplacer Essay, Reading Comprehension and Sentence Skills score recorded in the Fall 2001 suite of frozen extract files. This resulted in a sample of 371 English Composition students and 1,188 Developmental English students (see Table 1).

Design and Procedure

Student impact will be assessed in the following manner: (1) Identify cohort students who placed into English Composition and completed the course with a grade of C or better. Compare placement using the proposed cut scores with placement using actual placements. Would any of these students place into Developmental English under the proposed scores? (2) Identify first-time entering students who placed into English Composition and completed the course with a grade of C- or lower. Compare placement using the proposed cut scores with placement using actual placements. Would any of these students place into Developmental English under the proposed scores? (3) Identify cohort students who placed into Developmental English and completed the course with a grade of C or better. Compare placement using the proposed cut scores with placement using actual placements. Would any of these students place into English Composition under the proposed scores? (4) Identify first-time entering students who placed into Developmental English and completed the course with a grade of C- or lower. Compare placement using the proposed cut scores with placement using actual placements. Would any of these students place into English Composition under the proposed scores? (5) Use budgeted class size[1] to compute the number of sections to accommodate students using actual placements. (6) Use budgeted class size to compute the number sections to accommodate students using the proposed placements. (7) What are the implications for the colleges?

Definitions

College Size: Colleges are categorized as small, medium and large based upon average FTE generated.

Budgeted Class Size: Budgeted class size is used to estimate dollars needed to staff sections beyond those covered by full-time, general fund faculty and varies by college size and course type (see Table 2). Tables 3 and 4 illustrate the budgeted sections with actual sections offered.

Table 1: Population and Sample Identification (Fall 2001)

College	**# of New and Transfer Students Enrolled in Eng Composition**	**# of New and Transfer Eng Comp Students with all 3 Test Scores**
Asnuntuck	171	0
Capital	347	8
Gateway	605	12
Housatonic	554	75
Manchester	580	5
Middlesex	343	8
Naugatuck	536	0
Northwestern	185	93
Norwalk	663	5
Quinebaug	204	31
Three Rivers	494	5
Tunxis	375	129
System Total	**5057**	**371**
College	**# of New and Transfer Students Enrolled in Developmental English**	**# of New and Transfer Devl English Students with all 3 Test Scores**
Asnuntuck	132	0
Capital	421	5
Gateway	659	7
Housatonic	1197	500
Manchester	790	1
Middlesex	299	12
Naugatuck	789	15
Northwestern	276	114
Norwalk	750	107
Quinebaug	238	17
Three Rivers	206	0
Tunxis	807	410
System Total	**6564**	**1188**

---

[1] Class limits vary across the system, and at some colleges the class limits vary among sections of the same course. Because of this, budgeted class size was selected as the standardized unit of comparison.

Table 2: Budgeted Class Size

College	Size	Health	Bus & Mkt	Devl	Humanities	Math Sci	Other Occ	Social Sci	Technology	Uncoded
Asnuntuck	Small	15.00	21.56	18.75	21.56	21.56	21.56	21.56	15.00	21.56
Capital	Medium	15.00	22.81	18.75	22.81	22.81	22.81	22.81	15.00	22.81
Gateway	Medium	15.00	22.81	18.75	22.81	22.81	22.81	22.81	15.00	22.81
Housatonic	Medium	15.00	22.81	18.75	22.81	22.81	22.81	22.81	15.00	22.81
Manchester	Large	15.00	24.06	18.75	24.06	24.06	24.06	24.06	15.00	24.06
Middlesex	Medium	15.00	22.81	18.75	22.81	22.81	22.81	22.81	15.00	22.81
Naugatuck	Large	15.00	24.06	18.75	24.06	24.06	24.06	24.06	15.00	24.06
Northwestern	Small	15.00	21.56	18.75	21.56	21.56	21.56	21.56	15.00	21.56
Norwalk	Large	15.00	24.06	18.75	24.06	24.06	24.06	24.06	15.00	24.06
Quinebaug	Small	15.00	21.56	18.75	21.56	21.56	21.56	21.56	15.00	21.56
Three Rivers	Medium	15.00	22.81	18.75	22.81	22.81	22.81	22.81	15.00	22.81
Tunxis	Medium	15.00	22.81	18.75	22.81	22.81	22.81	22.81	15.00	22.81

Table 3: First College Level English Course Enrollments (Fall 2001)

College	Size	Sections	Enrollment	Limit	Credit Hrs	FTE	Budgeted Class Size	Estimated Sections	Difference
Asnuntuck	Small	7	171	125	3	34.20	21.56	7.93	0.93

Capital	Medium	15	347	345	3	69.40	22.81	15.21	0.21
Gateway	Medium	23	605	575	3	121.00	22.81	26.52	3.52
Housatonic	Medium	23	554	525	3	110.80	22.81	24.29	1.29
Manchester	Large	25	580	593	3	116.00	24.06	24.11	-0.89
Middlesex	Medium	14	343	350	3	68.60	22.81	15.04	1.04
Naugatuck	Large	24	536	560	3	107.20	24.06	22.28	-1.72
Northwestern	Small	7	185	174	3	37.00	21.56	8.58	1.58
Norwalk	Large	29	663	720	3	132.60	24.06	27.56	-1.44
Quinebaug	Small	9	204	216	3	40.80	21.56	9.46	0.46
Three Rivers	Medium	22	494	523	3	98.80	22.81	21.66	-0.34
Tunxis	Medium	15	375	375	3	75.00	22.81	16.44	1.44
System		213	5057	5081		1011.40		219.07	6.07

Table 4: Developmental English Course Enrollments (Fall 2001)

College	Size	Sections	Enrollment	Limit	Credit Hours	Student FTE	Budgeted Class Size	Estimated Sections	Difference
Asnuntuck	Small	6	132	175	3	26.40	18.75	7.04	1.04
Capital	Medium	23	421	478	3	84.20	18.75	22.45	-0.55
Gateway	Medium	27	659	655	3	131.80	18.75	35.15	8.15
Housatonic	Medium	52	1197	1204	3	239.40	18.75	63.84	11.84
Manchester	Large	38	790	836	3	158.00	18.75	42.13	4.13
Middlesex	Medium	15	299	485	3	59.80	18.75	15.95	0.95
Naugatuck	Large	37	789	806	3	157.80	18.75	42.08	5.08
Northwestern	Small	14	276	260	3	55.20	18.75	14.72	0.72
Norwalk	Large	36	750	752	3	150.00	18.75	40.00	4.00
Quinebaug	Small	12	238	420	3	47.60	18.75	12.69	0.69
Three Rivers	Medium	12	206	247	3	41.20	18.75	10.99	-1.01
Tunxis	Medium	36	807	829	3	161.40	18.75	43.04	7.04
System		308	6564	7147		1312.80		350.08	42.08

Results

English Composition – Student Impact

At the start of the semester 371 (7%) of the 5,057 students enrolled in English Composition met the criteria for inclusion in this study. Table 5 provides a breakdown of the grade and transcript notations of the 371 English Composition students. At the end of the semester, 200 (54%) of these students received a grade of C or higher, 105 (29%) received some other grade or transcript notation, and 66 (18%) students withdrew from their college (see Table 5).

Table 5: End of Term Grades – English Composition

	Frequency	Percent	Valid Percent	Cumulative Percent
A	43	11.59%	14.10%	14.10%
A-	28	7.55%	9.18%	23.28%
B+	34	9.16%	11.15%	34.43%
B	36	9.70%	11.80%	46.23%
B-	27	7.28%	8.85%	55.08%
C+	18	4.85%	5.90%	60.99%
C	14	3.77%	4.59%	65.58%
C-	12	3.23%	3.93%	69.51%
D+	3	0.81%	0.98%	70.49%
D	4	1.08%	1.31%	71.80%
D-	1	0.27%	0.33%	72.13%
F	31	8.36%	10.16%	82.30%
AW	3	0.81%	0.98%	83.28%
I	20	5.39%	6.56%	89.84%
IP	1	0.27%	0.33%	90.17%
M	1	0.27%	0.33%	90.49%
N	1	0.27%	0.33%	90.82%
NC	17	4.58%	5.57%	96.40%
W	10	2.70%	3.28%	99.67%
8 8	1	0.27%	0.33%	100.00%
Missing	66	17.79%		
Total	371	Valid Total	305	

If the proposed scores had been in place during the fall of 2001, the placements would have changed as shown in Table 6. Of the 200 students who completed English Composition with a grade of C or higher, 135 (68%) would have placed into Developmental English. Of the 105 students who received a grade of C- or lower or some other transcript notation, 73 (70%) would have placed into Developmental English. Of the 66 students who dropped out of their college, 40 (61%) would have placed into Developmental English. In all, of the 371 students in our English Composition sample, 248 (67%) would have placed into Developmental English.

Table 6: New Placements

		New Placement			
	ENG COMP	ENG COMP		DEVL ENG	
C or Higher	200	65	33%	135	68%
Other	105	32	30%	73	70%
Missing	66	26	39%	40	61%
Total	371	123	33%	248	67%

Table 7 provides a closer look at the 171 students in the "Other" and "Missing Categories" who would have placed in Developmental English under the proposed placement scores. Of the 51 students who received a C-, D+, D, D- or F in English Composition, 27 (53%) would have placed in Developmental English. Of the 44 students with some other transcript notation besides W, 37 (84%) would have placed in Developmental English. Of the 10 students who withdrew from the course (W grade designation), and not the college, nine (90%) would have placed into Developmental English. Of the 66 students who left their college, 40 (61%) would have placed in Developmental English. In total 113 (66%) of the 171 students in the "other" and "missing" categories may have been more appropriately placed into Developmental English as a result of the proposed placement scores, but there is no way of knowing for certain if the new placement would have been more accurate or if the student would have been any more successful.

Table 7: Detail of "Other" and "Missing"

	ENG COMP	DEVL ENG	PERCENT
C-	12	6	50.00%
D+	3	3	100.00%
D	4	2	50.00%
D-	1	0	0.00%
F	31	16	51.61%
AW	3	3	100.00%
I	20	13	65.00%
IP	1	1	100.00%
M	1	1	100.00%
N	1	1	100.00%
NC	17	17	100.00%
W	10	9	90.00%
8 8	1	1	100.00%
Missing	66	40	60.61%

If only standard letter grades of A through F are considered, 135 (68%) of the 200 students passing English Composition with a grade of C or higher would have placed in Developmental English if the proposed placement scores had been used. Twenty-seven (53%) of the 51 students receiving a standard letter grade of C- through F would have placed into Developmental English (see Table 8).

Table 8: New Placements (Grades A through F in English Composition)

		New Placement			
	ENG COMP	ENG COMP		DEVL ENG	
C or Higher	200	65	33%	135	68%
C- through F	51	24	47%	27	53%

Acknowledging known limitations of using grades as an outcome measure, the data indicate that 135 (54%) of the 251 English Composition students would have been delayed and 27 (10%) may have been more appropriately placed if the proposed placement scores had been applied.

Developmental English – Student Impact

At the start of the semester 1,188 (18%) of 6,564 Developmental English students met the criteria for inclusion in this study. At the end of the semester, 703 (59%) of these students received a grade of C or higher and 384 (32%) received a grade of C- or lower or some other transcript notation, while 101 (9%) students withdrew from their college[2] (see Table 9).

Table 9: End of Term Grades – Developmental English[3]

	Frequency	Percent	Valid	Cumulative Percent
A	91	7.66%	8.37%	9.15%
AO	2	0.17%	0.18%	9.33%
A-	78	6.57%	7.18%	16.51%
B+	91	7.66%	8.37%	24.88%
B	131	11.03%	12.05%	36.93%
BO	2	0.17%	0.18%	37.12%
B-	112	9.43%	10.30%	47.42%
C+	80	6.73%	7.36%	54.78%
C	115	9.68%	10.58%	65.36%
CO	1	0.08%	0.09%	65.45%
C-	61	5.13%	5.61%	71.06%
D+	11	0.93%	1.01%	72.08%
D	22	1.85%	2.02%	74.10%
D-	13	1.09%	1.20%	75.30%
F	91	7.66%	8.37%	83.67%
F*	5	0.46%	0.46%	84.13%
I	16	1.35%	1.47%	85.60%
IP	19	1.60%	1.75%	87.35%
M	32	2.69%	2.94%	90.29%
NA	13	1.09%	1.20%	91.49%
NC	72	6.06%	6.62%	98.11%
NS	3	0.25%	0.28%	98.39%
S	5	0.42%	0.46%	98.85%
U	1	0.08%	0.09%	98.94%
W	17	1.43%	1.56%	100.50%
X	3	0.25%	0.28%	100.78%
Missing	101	8.50%		
Total	1188	Valid Total	1087	

---

[2] "Missing" means withdrawal from a college while "W" indicates withdrawal from a course.

[3] Percents do not sum to 100 due to rounding errors.

If the proposed Accuplacer scores had been in place during the fall of 2001, all of the Developmental English placements would have remained the same. Because there is no change in placement, there is also no change in class section distribution.

Budgeted Sections/Section Distribution

There were 213 English Composition sections offered in Fall 2001, representing 4% of all general fund sections (see Table 3). Dividing enrollment by budgeted class size, results in an estimated 219 sections, six sections more than the colleges ran. There were 308 Developmental English sections offered in Fall 2001, which represents 6% of all general fund sections (see Table 4). Dividing enrollment by budgeted class size results in an estimated 350 sections, 42 sections more than the colleges ran. As shown in Table 10, the sample of 371 English Composition students accounted for 16.56 budgeted sections of English Composition based on the budgeted class size for humanities classes. Had the proposed placement scores been used, these same students would have accounted for 5.46 budgeted English Composition sections and 13.23 budgeted Developmental English sections.

Table 10: Budgeted Sections

		Humanities	Actual Placement (Eng Comp)		New Placements (Eng Comp)		New Placement (Devl Eng)	
College	**Size**	**Budgeted Class Size**	**Enroll.**	**Budgeted Sections**	**Enroll.**	**Budgeted Sections**	**Enroll.**	**Budgeted Sections**
Asnuntuck	Small	21.56	0	-	0	0.00	0	0.00
Capital	Medium	22.81	8	0.35	0	0.00	8	0.43
Gateway	Medium	22.81	12	0.53	4	0.18	8	0.43
Housatonic	Medium	22.81	75	3.29	0	0.00	75	4.00
Manchester	Large	24.06	5	0.21	1	0.04	4	0.21
Middlesex	Medium	22.81	8	0.35	0	0.00	8	0.43
Naugatuck	Large	24.06	0	-	0	0.00	0	0.00
Northwester	Small	21.56	93	4.31	16	0.74	77	4.11
Norwalk	Large	24.06	5	0.21	0	0.00	5	0.27
Quinebaug	Small	21.56	31	1.44	10	0.46	21	1.12
Three Rivers	Medium	22.81	5	0.22	2	0.09	3	0.16
Tunxis	Medium	22.81	129	5.66	90	3.95	39	2.08
System			371	16.56	123	5.46	248	13.23

Using the proposed scores to accommodate the initial cohort of English Composition students would have required an additional 2.13 budgeted class sections. This represents a 13% increase over the initial budgeted sections allocation.

Table 11 provides a system summary of the impact of using the proposed placement scores. The 371 English Composition students accounted for 16.56 budgeted English Composition sections. The 1,188 Developmental English students accounted for 63.36 budgeted Developmental English sections. The proposed cut scores would have no impact on the initial cohort of Developmental English students, but the distribution of the English Composition students shifted considerably. This shift would result in 123 students representing 5.46 budgeted

English Composition sections and 1,436 students representing 76.59 Developmental English sections. For this group, there is a 2.6% increase in the number of sections that would have to be offered to accommodate the same number of students under the proposed scores (82.05 proposed versus 79.92 actual).

Table 11: System Summary

Actual				With Proposed Placement Scores			
Eng Comp		Devl Eng		Eng Comp		Devl Eng	
Enrollment	Sections	Enrollment	Sections	Enrollment	Sections	Enrollment	Sections
371	16.56	1188	63.36	123	5.46	1436	76.59

Discussion

The results of this study show that a large number of successful English Composition completers would have placed into Developmental English using the proposed placement scores. One medium sized college, for example, had 75 students in the English Composition cohort. Under the proposed Accuplacer scores, all 75 would have placed into Developmental English, including 46 (61%) students who received a grade of C or higher in English Composition. Among the 200 successful English Composition completers, 135 would have placed into Developmental English. In other words 68% of the successful completers (grade C or higher) would have had to complete at least one Developmental English course and an English Composition course in order to attain the same outcome.

A total of 171 English Composition students received a grade of C- or lower, some other transcript notation, or withdrew from their college. One hundred thirteen of these students would have placed into Developmental English, but there is no way of knowing for certain if the new placement would have been more accurate or if the students would have been any more successful.

Regarding budgeted sections, the proposed scores would result in a reduction in English Composition enrollment by 248 students and a reduction of 11.10 budgeted English Composition sections while there would be a concomitant increase in Developmental English enrollment by 248 students and an increase of 13.23 budgeted Developmental English sections. The additional 2.13 sections is a 2.6% increase in the number of sections needed to accommodate the same number of students. This is a direct result of the budgeted class size difference between Developmental English and English Composition.

Can we generalize our findings back to the population to ascertain the overall impact to the system? The answer is no. Our sample is not random. The likelihood for selection based on the established criteria varies by college, and many students taking the Accuplacer are not taking all three English tests and/or their scores are not being entered into Banner at the same rate.

This study also illustrates that the condition and availability of data in Banner for in-depth analysis are not currently in place. One step towards improving the current situation is to put in place a common, standardized system-wide set of Accuplacer scores for placement into English Composition. However, the limited data available suggest that the currently proposed scores be reconsidered.

ENHANCING OUTCOMES ASSESSMENT BY DISCOVERING ALUMNAE/I SUCCESS STRATEGIES

Anne Marie Delaney
Director of Institutional Research
Babson College

Introduction

This paper presents an alumnae/i research study as a model for analyzing alumnae/i career development and enhancing the impact of higher education on alumnae/i career success. Primary goals of this study were to explore alumnae/i career paths; to examine the relationships between personal background characteristics, career experiences, opportunities, attitudes and personality characteristics and career success; and to enhance higher education's effectiveness in preparing graduates for career success. The major research questions of the study were:

- How do personal background characteristics affect career success?
- What effect do career strategies and experiences have on career success?
- Do certain attitudes and personality characteristics promote career success?
- How can higher education enhance graduates' potential for career success?

Review of the Literature

Alumnae/i research offered a rich resource of ideas for developing the conceptual framework and design of the present study. Schein (1974), Kotter (1995), Tharenou (1997) and Solymossy (1998) provided the major theoretical foundation for this research. Looking through different lenses, these authors identified individual and organization variables significantly related to career success.

Schein (1974), who conducted a 10-year follow-up study of graduates of MIT's Sloan School of Management, created the concept of career anchors to explain differences in career motivation and individual occupational aspirations. The five career anchors are managerial competence, technical/ functional competence, organizational security, creativity, and autonomy. Kotter (1995), who studied the career progress of graduates of the Harvard Business School, discovered new rules for achieving success, such as not relying on convention; monitoring globalization and its consequences; leading,in addition to just managing; and realizing that lifelong learning is increasingly necessary for success. Tharenou (1997) established the following career opportunities as predictors of success: successful performance in critical functional areas; experience on high visibility projects; employment in corporate headquarters; and breadth of experience in several functional areas. She also found individual traits -- ambition, intelligence, and high achievement motivation -- to be related to success.

Several studies documented significant relationships between personal attributes and career success. Cox and Harquail (1991) identified significant differences by gender.

Women experienced lower salary progression than men of comparable education, performance, age, and experience. Judge et al. (1999) determined that personality characteristics, including high extroversion and high conscientiousness, were significant predictors of career success over a 50 year span. Solymossy (1998) found that the economically successful exhibited higher levels of opportunistic behavior and of uncertainty, were more autonomous, showed higher levels of aggressive achievement, and were more goal-oriented. MacCrimmon and Wehrung (1990) determined that the most successful executives were the biggest risk-takers. Harrell and Alpert (1989) reported that social extroversion was the most valid personality predictor of success in terms of earnings, and Harrell (1969) documented that high earners were overwhelmingly in the socially desirable direction on personality measures.

Research Model

The research cited in the review of the literature served as the primary basis for developing the research model for the present study. As shown in Figure 1, the model consists of four groups of independent variables predicting success: personal background characteristics; career experiences; career opportunities and strategies; and attitudes/personality characteristics.

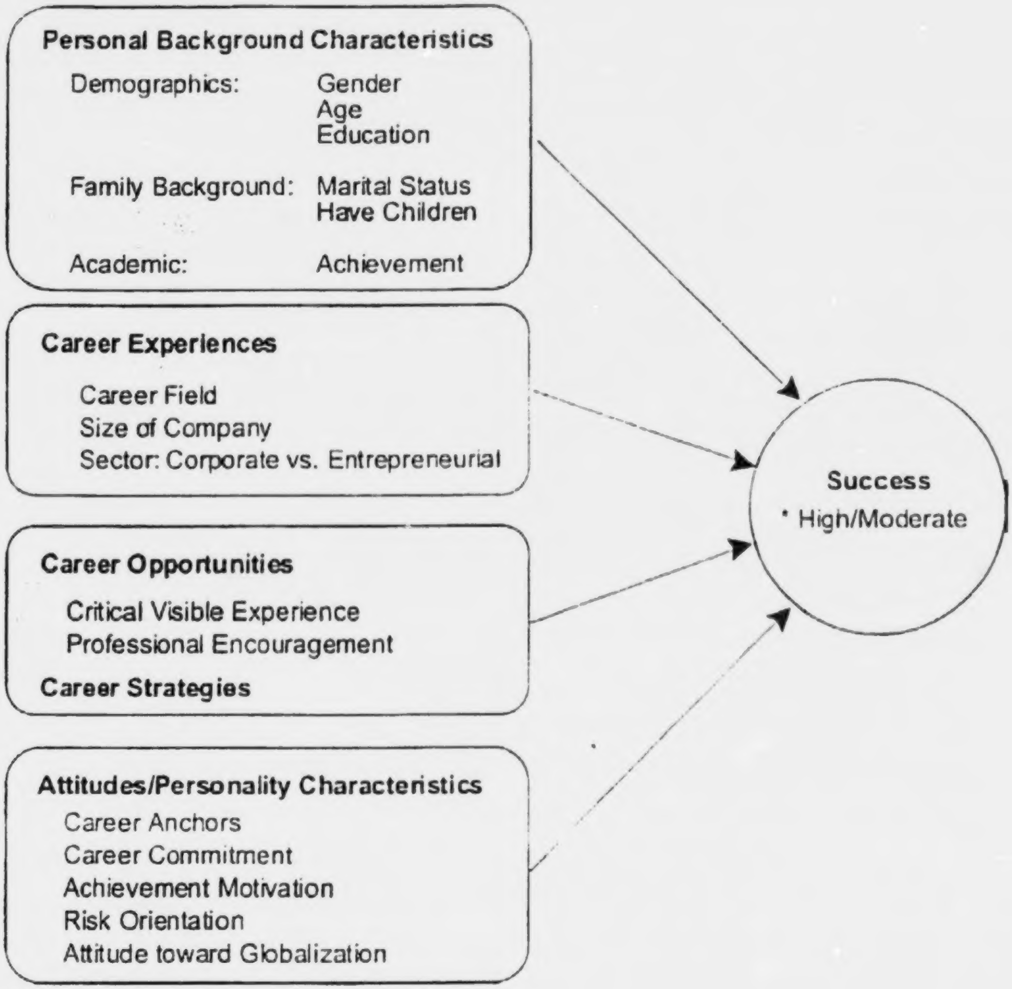
Personal Background Characteristics. As noted in the literature review, Gattiker and Larwood (1987) identified demographic variables as the best predictors of career success and mobility. The model for this study included the following demographic variables identified as significant predictors of success in previous studies: gender (Cox & Harquail, 1991; Tharenou, 1997); age (Gould & Penley, 1984; Gattiker & Larwood, 1987); marital status (Gattiker & Larwood, 1987; Tharenou, 1997); having children (Gattiker & Larwood, 1987); and education (Gould & Penley, 1984; Gattiker & Larwood, 1987; and Tharenou, 1997). Although previous studies (Harrell, 1969; and Harrell & Alpert, 1989) identified only second year MBA grades as significantly related to career success, this study included a more general measure of academic achievement, average grades in college, as a potential predictor of career success.

Career Experiences. This research model incorporated aspects of career experience previously determined to be related to career success: career field, size of company, and work sector - corporate versus entrepreneurial. Kotter (1995) found small size and entrepreneurial orientation to be related to career success.

Career Opportunities. Tharenou (1997) identified several career opportunities as significantly related to managerial advancement. Based on this research, the present model included two categories of career opportunity variables: critical visible experiences and professional encouragement.

Career Strategies. The model also incorporated career strategies as a potential predictor of career success. Results from previous research vary with regard to the influence of career strategies on career success. While Gattiker and Larwood (1987) found no relationship, Gould and Penley (1984) and Tharenou (1997) identified a significant relationship between career strategies and career success.

Figure 1: Predicting Alumnae/i Career Success


Personal Background Characteristics
Demographics:
Gender
Age
Education
Family Background:
Marital Status
Have Children
Academic:
Achievement
Career Experiences
Career Field
Size of Company
Sector: Corporate vs. Entrepreneurial
Career Opportunities
Critical Visible Experience
Professional Encouragement
Career Strategies
Attitudes/Personality Characteristics
Career Anchors
Career Commitment
Achievement Motivation
Risk Orientation
Attitude toward Globalization
Success
* High/Moderate

* High Success: Senior Management or: 1976 alumnae/i ≥ $200,000; 1996 MBA ≥ $125,000; 1996 Undergraduate ≥ $60,000

Attitudinal and Personality Characteristics. This segment of the model incorporates attitudes and personality characteristics found to be related to career success. The attitudinal constructs comprise: career anchors, career commitment, and attitude toward globalization. A five-item question was created to reflect Schein's (1974) five anchors: (1) managerial competence, (2) technical/functional competence, (3) organizational security, (4) creativity, and (5) autonomy. A slightly modified version of Carson and Bedeian's (1994) measure was used to determine career commitment and a series of statements were created to measure attitude toward globalization, which Kotter (1995) found to be related to career success. The specific items in these scales are described later in the methodology section. The two personality constructs in the model are: risk orientation (Kotter, 1995; Solymossy, 1998; MacCrimmon & Wehrung, 1990; and Sagie & Elizur, 1999), and achievement motivation (Kotter, 1995; Solymossy, 1998; and Tharenou, 1997).

Success. Study participants were classified into two groups: 'highly' and 'moderately' successful alumnae/i. Alumnae/i classified as 'highly' successful included those whose most recent positions were in senior management or who reported the following incomes for 1999: $200,000 or more for 1976 alumnae/i; $60,000 or more for 1996 undergraduate alumnae/i; and $125, 000 or more for 1996 MBA alumnae/i.

Data Source

Data were collected by means of a mailed questionnaire sent to Class of 1976 and 1996 undergraduate and graduate alumnae/i. Results are based on responses from 336 alumnae/i, including 49 Class of 1976 undergraduate alumnae/i, 85 Class of 1976 MBA alumnae/i, 101 Class of 1996 undergraduate alumnae/i, and 101 Class of 1996 MBA alumnae/i. Of the 336 study participants, 51 percent were classified as 'highly' successful.

Methods of Analysis

Bivariate techniques -- Chi-Square and t tests -- were used to determine if there were significant relationships between demographic, academic, and personality characteristics; career choices, career opportunities and success. Discriminant analysis was used to identify the most significant predictors of success. Analyses were conducted with individual items and with computed scales measuring career commitment, career opportunities experienced, and attitude toward globalization. Table 1A presents the names and statistical properties of the scales, and Table 1B identifies the items comprising each scale. As shown, the reliability coefficients range from .70 to .85.

Table 1A. Statistical Properties of the Scales

Scales	Mean	Standard Dev.	Reliability	No. of Items	Range of Responses
Career Commitment					
Personal Investment in Career	3.56	.89	.84	4	1-5
No Career Plan or Strategy	2.30	.78	.78	4	1-5
Perceived Career Costs	2.58	.99	.85	3	1-5
Career Opportunities					
Critical/Visible Experience	3.34	.55	.78	6	1-4
Professional Encouragement	3.29	.61	.70	2	1-4
Attitude Toward Globalization	3.68	.88	.80	3	1-5

Table 1B. Questionnaire Items Comprising the Scales

Career Commitment

Personal Investment in Career

- My career is an important part of who I am
- This career field has a great deal of personal meaning to me.
- I do not feel emotionally attached to this career field. *
- I strongly identify with my chosen career field.

No Career Plan or Strategy

- I do not have a strategy for achieving my goals in this career field.
- I have created a plan for my development in this career field. *
- I do not identify specific goals for my development in this career field.
- I do not often think about my personal development in this career field.

Perceived Career Costs

- Given the problems, I sometimes wonder if I get enough out of this career field.
- I sometimes wonder if the personal burden of this career field is worth it.
- The discomforts associated with my field sometimes seem too great.

Response Scale: 1 'Strongly Disagree' to 5 'Strongly Agree'

Career Opportunities

Critical Visible Experience

- Breadth of experience in several functional areas
- Broad management experience
- Experience in high visibility projects
- Experience in negotiating business deals
- Responsibility for major organizational assignments
- Successful performance on a critical project

Professional Encouragement

- Encouragement from colleagues
- Encouragement from supervisors

Response Scale: 1 'Not at All' to 4 'Very Much'

Attitude Toward Globalization

- Changes in the international economy have very little effect on my business. *
- Events in other parts of the world make very little difference in my business. *
- What is happening in the world economy seems remote from everyday business. *

Response Scale: 1 'Strongly Disagree' to 5 'Strongly Agree'

* These items were reverse coded for the creation of the scale.

Results

Results revealed statistically significant relationships between career success and personal background characteristics, career experiences, career opportunities and strategies, and attitudes/personality characteristics.

Differences by Personal Background Characteristics. Several statistically significant relationships were found between demographic characteristics and career success. Correlation analysis showed a small, statistically significant relationship between age and success ($r = .15, p \leq .05$), and a substantial, statistically significant difference was found between success and gender ($X^2 = 13.16, p \leq .001$). A significantly higher percent, 55 percent of the males, compared with only 30 percent of the females, were classified as 'highly' successful.

Marital status was also significantly related to success ($X^2 = 5.17, p \leq .05$). Some 55 percent of those currently or previously married, compared with only 41 percent of those who were single, were classified as 'highly' successful. Having children was also significantly related to success ($X^2 = 5.25, p \leq .05$). Fifty-seven percent of those who had children, compared with 44 percent of those who did not have children, were classified as 'highly' successful. Further analysis revealed no significant interaction effect by gender between marital status, having children and success.

No statistically significant relationships were found between educational attainment, average grade in college and career success. The success breakdown was approximately even among those earning a bachelor's and master's degree, and 51 percent of those earning A's or B's, compared with a slightly lower 44 percent among those earning an average grade of C, in college were classified as 'highly' successful.

Importance of Career Experiences. In terms of career experiences, results revealed statistically significant relationships between level of responsibility of first position ($X^2 = 16.71, p \leq .01$), career field ($X^2 = 28.21, p \leq .001$), size of company of most recent position ($X^2 = 15.85, p \leq .01$), and career success. Those who had senior level responsibility in their first position; who had entrepreneurial experience; and who recently worked for a small company were more likely to be classified as 'highly' successful.

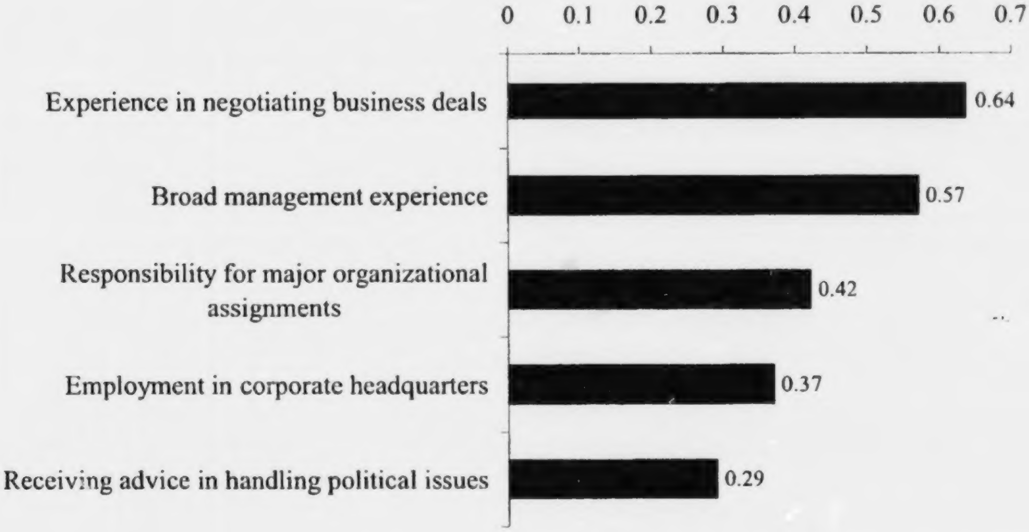
Advantage of Career Opportunities. As shown in Table 2, 'highly' and 'moderately' successful alumnae/i differed significantly on eight of the 13 career opportunities addressed in this study. Most notably, 'highly' successful alumnae/i reported more extensive experience in negotiating business deals, broad management experience, and responsibility for major organizational assignments. T test analysis identified a statistically significant relationship between alumnae/i scores on the *Critical Visible Experience* scale and career success ($t = 6.96, p \leq .001$), indicating the 'highly' successful had greater breadth of experience in several functional areas; broad management experience; experience in high visibility projects and in negotiating business deals; major organizational assignments; and successful performance on a critical project.

Table 2
Statistically Significant Differences in Career Opportunities Between 'Highly' and 'Moderately' Successful Alumnae/i

Career Opportunities	Highly Successful	Moderately Successful	Mean Difference	t Ratio
Experience in negotiating business deals	3.43	2.79	.64	6.16***
Broad management experience	3.48	2.91	.57	6.37***
Responsibility for major organizational assignments	3.43	3.01	.42	4.68***
Employment in corporate headquarters	3.23	2.86	.37	2.78**
Receiving advice in handling political issues	2.74	2.45	.29	2.45*
Breadth of experience in several functional areas	3.52	3.27	.25	2.94**
Experience in high visibility projects	3.58	3.33	.25	3.15**
Successful performance on a critical project	3.78	3.58	.20	3.32***

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Figure 2. Significant Mean Differences in Selected Career Opportunities Between 'Highly' and 'Moderately' Successful Alumnae/i



Significance of Attitudes and Personality Characteristics. Bivariate analyses revealed statistically significant relationships between emotional attachment to one's career, having a career strategy, and alumnae/i success status. 'Highly' successful alumnae/i were more likely to disagree with the following statements: "I do not feel emotionally attached to this career field" ($t = 2.14$, $p \leq .05$); and "I do not have a strategy for my development in this career field" ($t = 2.43$, $p \leq .05$). T test analysis also confirmed a statistically significant relationship between scores on the personal investment in career scale and success ($t = 2.03$, $p \leq .05$), indicating that that 'highly' successful alumnae/i consider their career as an important part of who they are. It has a great deal of personal meaning to them, and they feel emotionally attached and strongly identify with their career field.

Statistically significant differences were found between 'highly' and 'moderately' successful alumnae/i on two personality characteristics: achievement motivation and risk orientation. 'Highly' successful alumnae/i reported significantly stronger agreement with the statement, "I enjoy the uncertainty and risks of business" ($t = 4.37$, $p \leq .001$). In contrast, 'moderately' successful alumnae/i reported significantly stronger agreement with statements claiming they seldom get a sense of pride from their work ($t = -2.48$, $p \leq .05$); their goals are generally modest ($t = -3.50$, $p \leq .001$); and they need to know something has been done before they are willing to try it ($t = -2.83$, $p \leq .01$). Agreement with these statements indicates a lower level of achievement motivation and risk orientation.

The study also revealed that 'highly' successful alumnae/i were more aware of globalization, but had less time to attend cultural events. 'Highly' successful alumnae/i were significantly less likely to agree with the statement, "Changes in the international economy have very little affect on my business" ($t = -2.08$, $p \leq .05$). However, they were significantly more likely to agree with the statement, "The pressure of work doesn't leave much time to attend concerts or plays" ($t = 2.02$, $p \leq .05$).

Predicting Alumnae/i Success Status

Table 3 presents the results from the discriminant analysis employed to test a model predicting alumnae/i success status. The discriminant function coefficients indicate the relative weights for each variable found to be a significant predictor of alumnae/i success status. These results are based on analyses of 272 cases. This discriminant function model, including 11 variables, accurately predicted the success status of 78 percent of the 272 alumnae/i respondents. The canonical correlation of .51 indicates that this function explains 26 percent of the variance in success.

As shown in Table 3, gender is the single demographic characteristic included in the model. The negative coefficient indicates that males are more likely to be classified as 'highly' successful. Success is also determined by one's career experiences and strategies. Those who held higher levels of responsibility in their first position as well as those who recently worked for a smaller company were more likely to achieve success. Alumnae/i who developed and followed specific strategies were also more likely to achieve success.

These data confirm previous research findings linking certain career opportunities with success. Alumnae/i who had broad management experience, experience in negotiating business deals, and successful performance on a critical project were more likely to be classified as 'highly' successful. Personality characteristics that significantly predict success include risk orientation and achievement motivation. As indicated by the coefficient of .26, risk orientation is clearly the stronger predictor of success. Alumnae/i who report that they enjoy the uncertainty and risks of business are significantly more likely to be classified as 'highly' successful. Those whose goals and ambitions are not modest are also more likely to be classified as 'highly' successful.

These results also document relationships among cohort, satisfaction with first job, and success. Class of 1976 undergraduate alumnae/i and those who were satisfied with their first position were more likely to be 'highly' successful. The cohort effect may reflect the fact that this group had more time to achieve success, or that the income criteria used to define success may not have been comparable among the four cohorts.

Table 3
Discriminant Analysis Results: Predicting Alumnae/i Success Status

Predictors	**Standardized Discriminant Function Coefficients**	**Percent Correctly Classified**
Demographic Characteristics		78%
Gender	-.26	
Career Experiences		
Size of most recent company	-.24	
Level of responsibility of first job	.28	
Career Strategies	.05	
Career Opportunities		
Broad management experience	.28	
Experience in negotiating business deals	.29	
Success on critical project	.10	
Personality Characteristics		
Enjoy uncertainty and risks of business	.26	
Goals and ambitions not modest	.19	
Satisfaction with Career Experience		
Satisfaction with first position	.22	
Alumnae/i Class		
1976 Undergraduate Class	.27	
Canonical Correlation	.51	

$X^2 = 80.16$; 11 df; $p \leq .001$

In summary, these results from discriminant analysis indicate that alumnae/i, who are male; who held higher levels of responsibility in their first position; who eventually worked for a small company; who developed and followed career strategies; who had the following career opportunities: broad management experience, negotiating business deals, and successful performance on a critical project; who had high risk orientation and achievement motivation; and who were satisfied with their first position had a high potential for success in their careers.

Discussion

Personal Background Characteristics. Similar to findings of Cox and Harquail (1991) and Tharenou (1997), this study found a significant relationship between male gender and success. Also, similar to Gould and Penley (1984) and Gattiker and Larwood (1987), the study identified a significant correlation between age and success. Consistent with Gattiker and Larwood's (1987) and Tharenou's (1997) findings, this study identified a positive relationship between being married and achieving career success. Also similar to Gattiker and Larwood (1987), but in contrast to Tharenou (1997), this study documented a significant relationship between having children and success.

Results from this study revealed no statistically significant relationship between educational attainment and career success. In contrast, Gould and Penley (1984) and Tharenou (1997), respectively, had identified education as a significant predictor of salary progression and career advancement. With regard to academic achievement, this study identified no significant relationship between overall grades in college and career success. In their research, Harrell and Alpert (1989) found Harvard and Stanford MBA students' second year grades, but not first year grades or total GPA, to be correlated with earnings.

Career Experiences. Similar to Kotter (1995), this study documented significant, positive relationships among small size of company, entrepreneurial employment, and career success. Consistent with Schein's (1974) finding, the study also documented a significant relationship between success and career field, specifically managerial and entrepreneurial fields. Thus, the study confirmed the significance of the three career experience variables specified in the model: career field, size of company, and sector of employment. In addition, this study discovered that higher level of responsibility and satisfaction with one's first position correlated with success.

Career Opportunities. Findings from this study confirm Tharenou's (1997) research positing significant relationships between career success and the following career opportunities: negotiating business deals, broad management experience, responsibility for major organizational assignments, employment in corporate headquarters, receiving advice in handling political issues, breadth of experience in several functional areas, experience in high visibility projects, and successful performance on a critical project. In contrast, no significant relationships were found between success and assistance from informal social networks; beneficial training and development programs; encouragement from colleagues; encouragement from supervisors; and guidance from a mentor.

All but one of the experiences associated with success related directly to work; the one exception was receiving advice in handling political issues. In general, the interpersonal and social experiences did not relate significantly to the external measure of success used in this study. Further, of the two scales created to measure career opportunities, only the critical visible experience scale related significantly to career success. No significant relationship was found between the professional encouragement scale and career success.

Career Strategies. The majority of participants in this study reported they had developed and followed specific strategies to ensure their career success. While statistical analyses showed either a small or no statistically significant relationship between having a strategy and career success, this may be due to limitations of the measures employed in the study. Previously, Gould and Penley (1984) and Tharenou (1997) had identified specific career strategies as significantly related to success.

Attitudes/Personality Characteristics. Results from this study revealed a statistically significant relationship between career success and personal investment in one's career - one of the three dimensions in Carson and Bedeian's (1994) career commitment measure. However, no significant relationship was found between career success and the other two scales: no career plan or strategy and perceived career costs. Also, no significant relationship was found between success and the attitude statements designed to reflect Schein's (1974) career anchors. However, the absence of a relationship may be due to the limitations of the measure or the homogeneity of the study population.

Significant relationships found between high achievement motivation, risk orientation, and career success reconfirm findings from previous research (Kotter, 1995; Solymossy, 1998; and Tharenou, 1997), and the significant relationship between attitude toward globalization and success confirms Kotter's (1995) theory that a new rule of success is the ability to cope effectively with globalization.

Conclusion

Results from this study have significant implications for higher education. Findings identify personal attitudes, strategies, career experiences, and opportunities that enhance an individual's potential for success. For business students, these factors include high risk orientation, high achievement motivation, a positive attitude toward globalization, and opportunities to acquire broad management experience, negotiate business deals, and achieve success on a critical project. Based on these findings, the College should design programs to develop attitudes correlated with success and encourage students to seek the types of career opportunities found to enhance success. Thus, the study demonstrates how alumnae/i research may be used to enhance higher education's effectiveness in preparing students for success in their career. Finally, the research model provides a framework for designing alumnae/i studies and identifying predictors of success among graduates in other professional fields.

References

Carson, K. D., & Bedeian, A. G. (1994). Career commitment: construction of a measure and examination of its psychometric properties. Journal of Vocational Behavior, 44, 237-262.

Cox, T. H., & Harquail, C. V. (1991). Career paths and career success in the early career stages of male and female MBAs. Journal of Vocational Behavior 39, 54-75.

Gattiker, U. E., & Larwood, L. (1987). Career success, mobility and extrinsic career satisfaction: studying corporate managers. Paper presented at the Annual Meeting of the American Sociological Association, Chicago, Illinois.

Gould, S., & Penley. L. E. (l984). Career strategies and salary progression: A study of their relationships in a municipal bureaucracy. Organizational Behavior and Human Performance, 34, 244-265.

Harrell, T. W. (1969). The personality of high earning MBA's in big business. Personnel Psychology 22, 457-463.

Harrell, T. W., & Alpert, B. (1989). Attributes of successful MBAs: a 20-year longitudinal study. Human Performance 2(4), 301-322.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. Personnel Psychology 52, 621-652.

Kotter, J. P. (1995). The new rules: how to succeed in today's post-corporate world. New York: Free Press.

MacCrimmon, K. R., & Wehrung, D. A. (1990). Characteristics of risk-taking executives. Management Science 36(4), 422-435.

Sagie, A., & Elizur, D. (1999). Achievement motive and entrepreneurial orientation: a structural analysis. Journal of Organizational Behavior 20, 375-387.

Schein, E. H. (1974). Career anchors and career paths: a panel study of management school graduates (Technical Report No. 1). Cambridge, Massachusetts: Massachusetts Institute of Technology, Sloan School of Management.

Solymossy, E. (1998). Entrepreneurial dimensions: the relationship of individual, venture, and environmental factors to success (Doctoral Dissertation, Case Western University, 1998). Dissertation Abstracts International, 59, 5-A.

Tharenou, P. (1997). Managerial career advancement. International Review of Industrial and Organizational Psychology 12, 39-93.

DETERMINANTS OF STUDENT DROPOUT IN CRITICAL PERIODS: COHORT DIFFERENCES AT A VIRTUAL UNIVERSITY

Tae Young Han
Research Associate
&
Mitchell S. Nesler
Associate Dean for Outcomes Assessment and Research
Excelsior College

Abstract

By investigating a set of academic variables regarding student retention at a virtual university, this study found interaction effects on the relationship between graduation status and period of retention based on certain academic variables. The results provide insight into the utilization of timely interventions for students possessing specific characteristics.

Introduction

One of the biggest challenges facing non-traditional higher education is student retention. Research on student retention generally takes one of several theoretical approaches, demonstrating psychological, organizational, interactionist, sociological, or economic points of view (Kroc & Hanson, 2001). Among these perspectives, many theories pertaining to student retention were based on the interactionist model which focuses on the students' interaction with the college, especially with other students and faculty. Although various models within this approach stress different aspects (i.e., Bean, 1983; Tinto, 1987), they often overlap and are complementary (DesJardins, Ahlburg, & McCall, 1999). Furthermore, given that student retention or departure is a dynamic process, Ishitani & DesJardins (2002) added another dimension that other models lack in terms of practical application. Specifically, they emphasized the longitudinal nature of the process of student departure, suggesting that the dimension of time be considered in the process.

Besides adding a time dimension, studies regarding the influence of academic and social interactions in traditional higher education may need to be modified and applied differently in virtual university contexts. For example, students who make up this type of learning community are often older than students of traditional colleges. In addition, they tend to be full-time participants in the workforce, often managing the priorities of work and family life. As a result of various and conflicting environmental factors, educational pursuits are often the first endeavor to give up. To meet this challenge, virtual universities make an effort to assist students who are academically, geographically, technologically, and personally diverse in reaching their goals by offering a full range of flexible academic options. However, graduation rates for non-traditional students remain lower than their traditional counterparts (Horn & Carroll, 1996).

Considering more dynamic processes of student retention in distance learning, studies of retention trends or path models may be insufficient for institutional research in the sense that they do not specify when the retention determinants have a salient effect (i.e., Middaugh, 1992). As DesJardins et al. (1999) suggested, the information pertaining to the critical period of determinants may assist in exerting timely and intensive interventions for vulnerable students. For example, at traditional institutions, academic variables such as high school achievement or SAT scores have been reported to have significant impact on student success in early on. However, the extent of these influences declines as students progress through college (Kroc & Hanson, 2001). Thus, colleges may not expect student success later in the enrollment term simply based on students' high SAT scores. In similar way, once the extent of the influence of student factors at a particular period of enrollment is known, colleges can become more efficient in the use of interventions by identifying the students who need special care at a particular time.

Following the conceptual framework of event history modeling suggested by DesJardins et al., this study examines the influence of different factors across student cohorts in an effort to identify the pattern of influence that hinders or promotes student retention at certain times in their higher education career. Thus, the primary purpose of this study is to provide preliminary insight into the conditions by which proactive efforts by colleges can have optimum impact. Specifically, this study explores some critical factors that affect students' enrollment status, specifically graduation and withdrawal. It also looks at the influence of these factors at different points in time to see when their influences are strong across cohorts. "Cohort" in this study is defined as the students who graduated/withdrew at the same time after enrollment. Although some analyses are based on yearly cohorts (first through sixth year), most are based on raw information from the enrollment period (i.e., based on monthly degree conferrals) reflecting the graduation system of the virtual university.

In terms of predictive factors, several cognitive, non-cognitive, and demographic variables have long been identified as the important individual determinants of student retention in higher education (e.g., Bean & Eaton, 2001; Clagett, 1992). This study explores the effect of basic demographic variables on retention. This exploration may provide not only the pattern of retention by demographics, but also the sense of whether the main analysis should consider the interaction effect of the demographic variables.

Some academic background factors were also investigated. Since students of non-traditional institutions often have some college experience prior to enrollment, previous studies (e.g., Cabrera, Nora, & Castaneda, 1993) tested the effect of early year experience in college. This study will also look ate prior college experience, because students in the virtual university often come back to school long after their stopout.

Methodology

Sample

The sample consists of students who had enrolled and exited the institution (via graduation or withdrawal) between fiscal years 1993 and 1998 at a virtual university located in the Northeast. Graduate and withdrawn students from a school in the college were the subject of the analyses. Currently enrolled students were excluded. A total of

2712 students were initially included in the study; 1363 who withdrew, 1349 who graduated.

Demographic variables examined include gender, ethnicity, highest level of education attained, and military experience at the time of enrollment. Males (63.2%) and those who did not have military experience (81%) represented a major potion of the sample. Caucasians (65.4%) were also the majority, followed by African Americans (12.4%), Asians (8.6%), and Latinos (7.2%). In terms of highest educational level attained, individuals who held an associate degree (44.9%) and high school graduates (30.2%) made up the largest subgroups, followed by those with a certificate/diploma (10.2%) or baccalaureate degree (7.4%).

Procedure

A series of demographic and academic variables were examined in an attempt to identify which factors were related to retention rate. Although the main analyses were done using raw information from the enrollment period as the outcome variable, students were also grouped by cohorts to develop graphical patterns of the predictor variables. Students who enrolled in different years but stayed the same length of time (i.e., second year, third year,…, and sixth year) were merged into the same cohort. In each cohort, students were again categorized by their academic outcome status (i.e., graduation or withdrawal). Therefore, students with the same status in the same cohort could be extracted from the data of different enrollment years.

Results

Demographic characteristics

First, the graduation rate was examined in relation to selected demographic variables across the cohort to see if the retention pattern differed across groups. The retention patterns were not much different when viewed in terms of gender and military experience, ethnicity, and highest educational level. Although the graphics illustrate deviations at "certificate degree" group and "Asian" in 4-5 year cohort group (Figure 1 & Figure 2), chi-squared values for the cells were not significant. As a result, these patterns were not expected to be a significant influence on the main analyses.

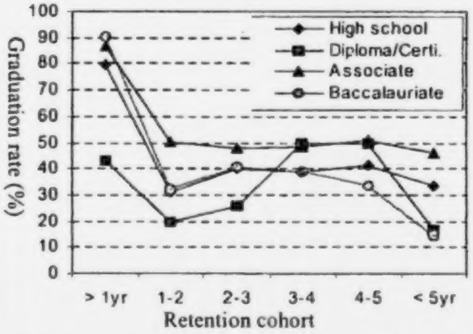


Figure 1. Retained period by highest education

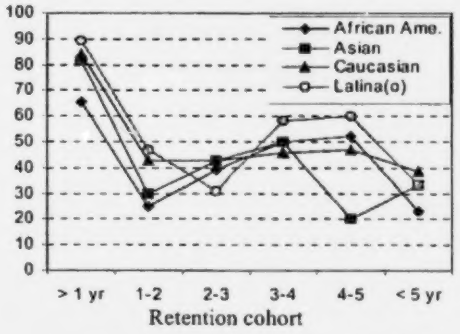


Figure 2. Retained period by ethnicity

Preliminary analyses of academic factors

Academic variables prior to student enrollment include the number of institutions previously attended, number of credits at initial status report, GPA at initial status report, and whether or not students used any exams for credit. Since the sample institution's evaluation system is based on credits taken from other institutions, the number of institutions reflects variability of the student credit background. T-tests revealed that students who graduated had higher GPAs, more initial credits, and took credits from more colleges than the withdrawn students, while they were retained for shorter periods in the sample school (Table 1).

To investigate the matter further, the relative strength of the variables on retention outcome status (i.e., graduate or withdrew) was analyzed. Using logistic regression analysis ($\chi 2$=525.7, p< .01), initial GPA was found to have the strongest impact on graduate/ withdrawal outcome, which is similar to that reported in previous studies (Table 2). The odds of graduating were nearly 65% higher than the odds of withdrawing outcome in terms of initial GPA, taking into account other factors in the analysis. The next strongest influence was the number of colleges attended prior to enrollment in the sample school.

Table 1. T-test of academic variables by retention outcome

Variables	Status	N	Mean	Std. Dev.	t
Retained period	Graduate	1349	1.84	1.51	** -4.24
	Withdrawn	1363	2.09	1.52	
Total Institutions	Graduate	1349	5.64	3.53	** 23.45
	Withdrawn	1353	2.72	2.92	
Initial GPA	Graduate	1026	3.10	0.47	** 6.95
	Withdrawn	719	2.93	0.52	
Initial credit	Graduate	1003	104.65	41.03	** 12.88
	Withdrawn	736	79.25	40.09	

Note. ** : p < .01, Sample size varies due to missing value.

By inverting the odd ratio calculation (i.e., Menard, 1996; DesJardins, 2001), the retained period was the third strongest influence on graduate/withdrawn distinction, with the odds of graduating being 30% lower than odds of withdrawing. The weak effects of initial credit and exam credit were analyzed more specifically in the later part of this study.

Table 2. Logistic regression analysis of academic variables on graduation status

Variable	b	Wald $\chi 2$	Odd Ratio	IOR	C.I of Odd Ratio	
Initial credit	.01	** 58.26	1.01		1.01,	1.01
Initial GPA	.50	** 19.09	1.65		1.32,	2.06
# of institution	.37	** 181.33	1.45		1.38,	1.52
Retained period	-.26	** 53.17	.77	1.30	.72,	.83
Exam credit	-.01	.01	.99	1.01	.99,	.99
Constant	-3.47	** 81.34				

**: p < .01, IOR: Inverted Odd Ratio, Exam credit is dichotomous by having or not having.

Student academics before enrollment

Among the variables that can be utilized to predict cohort differences prior to students' graduation or withdrawal, the initial credit hours and the initial GPA are variables that students bring to school at the point of enrollment. Because these variables do not involve interaction with the sample school, they are regarded as student factors that are external to the institution.

A. Initial credit hours: The students of the sample virtual university already had some college-level educational experience. Credits obtained prior to enrollment at the sample institution can be transferred and applied to the desired degree program, where appropriate. Although it is natural to think that students who brought more credits would remain enrolled for a shorter period of time in the institution, this is not always true because the credits students brought may or may not satisfy the requirements for the desired degree program. Further analyses were conducted to see if there is a difference between graduates and the withdrawn students in terms of initial credits and period retained relationships.

First, although the graduates and the withdrawn students were significantly different in terms of initial credit hours (M=105 vs. 79, t=12.9, $p < .01$, see Table 1), correlation analysis separating the two groups revealed that the relationship between initial credit hours and period retained is statistically significant only for the withdrawn group ($r = -.05$, n/s vs. .14, $p<.01$, respectively), implying an interaction effect with academic status. Thus, the interaction effect of status was examined using dummy coding in a multiple regression analysis. The interaction term was statistically significant, increasing R^2 by 0.8 percent ($B = .01$, $p < .01$, $R^2 = .018$, $p <.01$). Specifically, the more initial credits withdrawn students report, the shorter time they stayed at school (Table 3).

Table 3. Regression analysis of initial credit hours and status on retained period

Model	Variable	b	β	T	C.I. of b	
1	Intercept	2.41		** 38.70	2.28,	2.53
	Initial credit	-.002	-.05	-1.79	-.004,	.00
	Status	-.39	-.11	** -4.32	-.52,	-.20
2	Intercept	2.47		** 38.43	2.34,	2.59
	Initial credit (C)	-.01	-.15	** -3.95	-.01,	.00
	Status (S)	-.39	-.12	** -4.68	-.55,	-.23
	C * S	.01	.14	** 3.64	.00,	.01

Note: *: p<.05, **: p<.01, C.I.: Confidence interval, Model 1 uses main effect term only, Model 2 adds interaction term, Initial credit hour is mean-centered.

B. Initial GPA: Given the fact that a fair amount of time has passed since students completed high school education, traditional academic records (e.g., SAT score) may be a distal indicator regarding the student retention outcome. Rather, as most of the students in this study already have an experience in higher education, the GPA prior to enrollment may be comparable to the student's prior record in traditional higher-educational institutions.

The initial t-test between graduates and withdrawn students revealed that graduates brought higher initial GPAs than the withdrawn students when they enrolled. ($t = 6.95$, $p < .01$). In contrast, regression analysis indicated that the initial GPA did not predict how long students are retained. In addition, there was no interaction effect of the graduate status on initial GPA – retained period relationships. After looking at the relationship in detail, however, it was found that the relationship between initial GPA and the retained period was curvilinear (Figure 3). Graduates showed a u-shape relationship between initial GPA and stayed period, whereas the withdrawn students indicated inverted u-shape relationships.

Thus, the relationship for the graduates were better fit with the quadratic function,

$$\text{Period}_{(grad)} = 8.85 - 4.47(\textit{initial GPA}) + .72\,(\textit{initial GPA})^2$$

while the appropriate function for the withdrawn students was,

$$\text{Period}_{(with)} = -1.94 + 2.89(\textit{initial GPA}) - .46(\textit{initial GPA})^2$$

Adding the second-order squared term of initial GPA in the regression model, the graduation status was found to moderate the effect of the squared initial GPA on the retained period ($R^2 = .023$, $p < .01$).

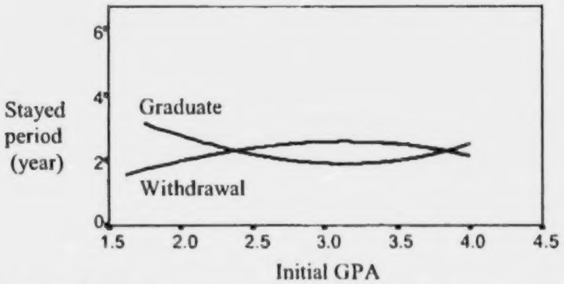


Figure 3. Effect of the status on the Initial GPA – Stayed period Relationships

The graduates revealed a negative GPA-retained period relationship when their initial GPA is lower than average, whereas the withdrawn students showed a positive GPA- retained period relationship. This pattern of interaction by graduate status was reversed for those who had a higher initial GPA (Table 4). The inflection point estimate was about 3.12 at their initial GPA. This result implies that students who report low initial GPAs should be paid more attention at the early period of enrollment because they are far more likely to withdraw soon after enrollment. Once students report about average initial GPA, withdrawn students were retained for a longer period of time than graduates, suggesting that schools need to pay more attention on medium range of initial GPA when they are at the later cohort group.

Table 4. Regression analysis for the initial GPA – stayed period relationship

Variable	b	β	t	C.I. of b
Intercept	2.45		** 20.78	2.22, 2.68
Initial GPA (G)	. 06	.02	.53	-.17, .30
Initial GPA^2(G2)	-.45	-.08	* -2.32	-.84, -.07
Status (S)	-.68	-.20	** -6.20	-.89, -.46
G * S	-.17	-.04	-1.04	-.49, .15
G2 * S	1.19	.16	** 4.18	.63, 1.74

Note: *: p<.05, **: p<.01

The effect of academic experience at the institution

The effect of academic experience (i.e., exam-taking experience and total number of institutions attended) on stayed period was also investigated to determine if this relationship is moderated by the graduation status. Although students who stay in school for a longer period of time are more likely to have an opportunity to take exams or to experience more institutions, it is assumed that the phenomena would differ by graduation status. One of these assumptions (the number of institutions attended) failed to prove a different effect by graduation status. The t-test in Table 1 may be enough to refer its effect.

Whether or not students took exams for credit (the other academic experience variable included in the analysis) interacted with the graduation status on retention period. Figure 4 indicates that although students who took exams were retained for a longer period in general, the difference is more salient for withdrawn students. Given the result that student status also interacted with initial credit hours, the relationships among the variables discussed here were analyzed using a regression model taking this aspect into account.

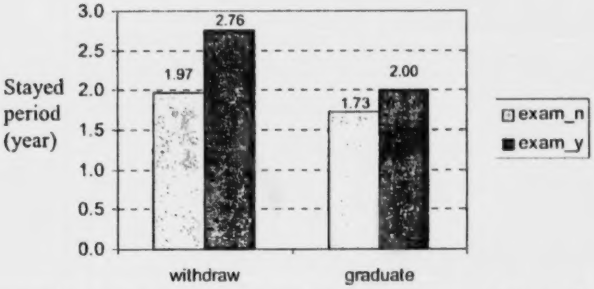


Figure 4. Comparison of exam-taken and not-taken group by graduation status

Table 5. Regression analysis of the effect of exam taken

Model	Variable	b	β	T	C.I. of b	
1	Intercept	2.33		** 36.19	2.20	2.46
	Status	-.441	-.13	** -5.19	-.61,	-.27
	Exam taken	.367	.11	** 4.30	.20,	.53
	Initial credit	-.002	-.04	-1.78	-.004,	.00
2	Intercept	2.32		** 32.44	2.18,	2.46
	Status (S)	-.343	-.10	** -3.41	-.54,	-.15
	Exam taken (E)	.668	.19	** 4.51	.38,	.96
	Initial credit (C)	-.005	-.14	** -3.64	-.01,	.00
	S * E	-.444	-.12	* -2.43	-.80,	-.09
	S * E * C	.002	.03	.99	.00,	.01

Note: *: p<.05, **: p<.01, Model 1 uses main effect term only, Model 2 adds interaction term,

When the three predictive variables (student status, exam taken experience, and initial credit hours) as well as their interaction terms regressed on the outcome variable, or retained period, the effect of exams was still moderated by student status ($R^2 = .022$, p <.01). This suggests the presence of motivational effects of the exam experience. In other words, once students take an exam, they are more likely to persist in their studies, which prolongs student retention. Thus, colleges may need to encourage students to take exams to acquire credit, or, at least, the information of who did not take exams yet can be utilized as additional information for retention management.

Discussion

This study primarily examined when student demographic and academic variables make a substantial difference on departure outcomes at a virtual university. The rationale of this approach was to empower higher-educational institutions to positively affect student outcomes and retention. If colleges have an idea of what variables produce meaningful differences between graduate and withdrawal outcomes as well as when these variables are most influential, retention efforts can be exerted more efficiently at a particular time focusing on the students who possess these characteristics. In contrast to research that explores simple retention-withdraw difference, this study provides a deeper understanding of retention patterns over time.

Using a sample from a virtual university from the time of enrollment to departure, this study in part supported the findings of previous studies, but also identified interaction effects. This study investigated the effect of a set of variables on the retained period. The academic variables students possessed prior to the time of enrollment, such as initial credits, initial GPA, and total number of institutions student attended, were analyzed. As an analytical strategy, this study primarily explored the interaction effect of student status on the predictor-retained period relationships, specifying linear or curvilinear interaction effects of variables on the retention outcome.

The interaction effect of initial credit found in this study has two implications in terms of student retention management at virtual universities. First, students who report fewer initial credit hours are more likely to withdraw. This may be a common phenomenon given that those who report fewer credits may need more time until graduation, during which several factors may influence persistence. Secondly, (an important implication for the institutional research) those who report far fewer credit hours are more likely to withdraw when enrolled for a longer time, while students who reported more credit hours do not show a significant difference in how long they are retained. This suggests that colleges need to focus more on students with minimal initial credits at their later academic years, because they are at risk of dropping out. In contrast, once students report more credits than average, they maintained enrollment for the same period of time regardless of their graduation status (i.e., graduate or withdraw).

The curvilinear relationship of initial GPA and enrollment period was also an interesting finding. Those students with lower initial GPA's are likely to either withdraw earlier, or persist longer if they graduate. On the contrary, students who reported moderate initial GPA's remained enrolled longer if they decide to withdraw, and for shorter periods if they graduate. In other words, even though graduates reported significantly higher initial GPA's than withdrawn students in general, the withdrawn students with lower GPA's were retained for shorter periods than those who withdrew with average GPA's, whereas graduates with lower GPAs maintained enrollment over a longer period than graduates with average GPA's. Thus, colleges may need to give more attention to the students with lower initial GPA's at the early stages of their enrollment. In contrast, students with moderate GPA's should be given more attention in the later part of the enrollment period because they are likely to withdraw later than the graduates with the similar GPA's.

In terms of the student exam taken experience at the college, students who have taken an exam generally appear to stay for a longer time. Moreover, students who have taken exams and persist longer than average are more likely to withdraw in the long run. Thus, institutions need to pay attention to students who have taken exams among the later cohort groups.

This study has some limitations. First, the effect sizes of the findings are relatively small, although they are highly significant. Given that there are several variables that account for the retention outcome (Kroc & Hanson, 2001), the limited number of variables this study dealt with may not be sufficient to explain retention phenomena. Furthermore, students in a virtual university are more diverse than those in traditional universities, which makes it difficult to explain variance for any single variable. However, since it is hard to find any dominant factor that accounts for student retention, even small variances may deserve notice in such novel contexts. For more practical application, some latent factors that are influencential in some smaller number of categories need to be explored. In addition, the longitudinal relationship this study investigated does not suggest a causal relationship between the variables. From the interactionist perspective, students may leave the school early due to poor interactions, *or* their interaction with (or experience in) the institution may be scant because they remain enrolled for relatively a short time. To examine if the effect of the predictive variables in this study is true causality, more sophisticated research design may necessary.

References

Bean, J. P. (1983). The application of a model of turnover in work organization to the student attrition process. *Review of higher education*, 6, 129-148.

Bean, J. P. & Eaton, S. B. (2001). The psychology underlying successful retention practices. *Journal of college student retention*, 3, 73-90.

Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *Journal of higher education*, 63, 143-164.

Clagett, C. (1992). Enrollment Management. In M. A. Whiteley, J. D. Porter, & R. H. Fenske (Eds.), *The primer for institutional research.* (pp. 12-24). Tallahassee, FL: Association for Institutional Research.

DesJardins, S. L. (2001). A comment on interpreting odds-ratios when logistic regression coefficients are negative, *AIR professional file,* 81, Fall.

DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event history model of student departure, *Economics of education review,* 18, 375-390.

Horn, L. J., & Carroll, C. D. (1996). *Nontraditional undergraduates: Trends in enrollment from 1986 to 1992 and persistence and attainment among 1989-1990 beginning postsecondary students* (Report No. NCES 97-578). U.S. Department of Education, Office of Educational Research and Improvement. [On-line] Available: www.ed.gov/NCES/pubs/97578.html.

Ishitani, T. T., & DesJardins, S. L (2002). A longitudinal investigation of dropout from college in the United States, *College student retention,* 4, 375-390

Kroc, R. & Hanson G. (2001). Enrollment management and student affairs. In R. D. Howard (Ed.), *Institutional research: Decision support in higher education.* (pp. 1- 59), Tallshassee, FL: Association for Institutional Research.

Menard, S. (2001). *Applied logistic regression analysis*. Sage publications, Thousand oaks, CA.

Middaugh, M. F. (1992). Persistence. In M. A. Whiteley, J. D. Porter, & R. H. Fenske (Eds.), *The primer for institutional research.* (pp. 1-11), Tallahassee, FL: Association for Institutional Research.

Tinto, V. (1987). *Leaving college; Rethinking the causes and cures of student attrition.* The University of Chicago Press, Chicago, IL.

LINKING INTERNAL TRANSFER PATTERNS TO COLLEGE STUDENT EXPERIENCES: A CASE STUDY

Jean O. Marriott
Research & Planning Analyst
Planning, Marketing, & Assessment
Carroll Community College

Catherine J. Alvord
Research & Planning Associate
Institutional Research & Planning
Cornell University

Abstract

At a large Research I University, the authors tracked students transferring among different schools of the institution. The paper will examine the connection between the internal transfer patterns, student outcome measures, and student responses on internal surveys. The analysis will reveal whether or not academic policy should be altered to affect existing internal transfer patterns.

Introduction

The Research I University studied encompasses eight undergraduate schools. Each unit houses its own majors of study, with minimal overlap among the schools at the undergraduate level. The university experiences a substantial volume of internal transfers, i.e., students transferring from one school to another within the same university. For the purposes of this paper, the authors concentrated on the group of outgoing internal transfers, i.e., those students transferring out of one of the eight undergraduate units into a different undergraduate unit.

Overall, the number of internal transfers has steadily decreased from 337 students to 279 over the past 5 years, a drop of about 17% in internal transfers for the fall semester. Although internal transfers have been more volatile for the spring semester, the university experienced a decrease in students internally transferring, from 272 to 235 students over the past 5 years, a drop of about 14%. Despite the overall decline in the number of internal transfers, there are some schools within the university that lose students to other units. The authors theorized that minimal movement between schools would indicate that the marketing programs of the schools are successful at targeting students who fit the schools' programs. Conversely, if there are schools that consistently lose their students to other units, then it is speculated that the effort to match students for certain majors of study could be improved.

Historically, persistence rates for the university as a whole remained constant over the five years under investigation. While the marketing plans successfully produced undergraduates that were a good match with the university, several individual schools had a substantial number of matriculants who were not reliable persisters within their own units. This raises the question of whether there are flaws in the marketing of those schools, and more specifically, whether students anticipate a different college experience, such as regimen of coursework, interactions with faculty and administration, or simply school culture? Are the internal transfer students less satisfied with their school?

Consistent student attrition in several schools within the university poses challenges on several fronts.

1. **Enrollment management within the schools becomes problematic.** The eight schools follow a strict enrollment plan recently adopted by the university. The objectives of the plan are to manage longer-term, university-wide, undergraduate enrollment targets for fall first-time freshmen; to protect individual schools from under-enrollment; and to discourage schools from over-enrollment in order to optimally maximize university resources without taxing them (Alvord, 2001). Instituting disciplined enrollment targets for only fall first-time freshmen stems from the fact that the university has the most control over this group, i.e., admissions committees can offer acceptance to a limited number of freshmen for each college.

2. Total undergraduate enrollment over an operational maximum taxes the physical and human resources of the university. In addition, significant year-to-year fluctuations create difficulties for both the teaching and infrastructure resources. To achieve the total enrollment goal, each college must successfully model all groups of students admitted to the units, including internal transfers. Unfortunately, the internal movement is difficult to predict from year to year; currently, the schools have reduced control over student transfers from other internal units, as compared to admittance of first-time freshmen. Thus, the complexity in forecasting the number of internal transfers creates challenges in enrollment modeling for the university overall.

   Instead of determining the precise number of internal transfers, the authors hypothesize that the university may improve enrollment modeling by limiting the number of transfers among each college. The authors surmise that the most reasonable method of decreasing movement of transfers is to match prospective students with the school that most closely fits their interests and needs. If first-time freshmen enter the school, and thus a program of study that most closely aligns with their interests, the necessity to establish the number of internal transfers would be greatly diminished and would perhaps become unnecessary. Tracking transfers and the reasons that precipitate them will offer the schools some insight so that administration may adjust academic policy to make it more seamless for future internal movement. Thus, it will enhance enrollment management, since the ability to internally transfer within the university is a distinctive strength.

3. **Financial aid models may potentially be affected.** The university awards financial aid based on need. Approximately 85% of students receive need-based aid on a percentage instead on a dollar basis. At this land grant university there are two separate tuition based units; endowed for the private sector and in-state and out-of-state for the contract sector. Therefore, students transferring from a lower tuition

based unit to a higher one potentially increase their need for financial aid resources. Conversely, students transferring from a higher tuition based unit to a lower one would decrease the need for institutional financial aid resources.

4. **Time to degree may be affected.** Students changing schools within the institution may take longer to graduate.
5. **Increase in attrition rate may occur.** The student is more likely to withdraw altogether from the institution prior to degree completion if the student is poorly integrated into the school of initial matriculants, given that most attrition happens early in a student's post-secondary career. If students believe that they are not academically integrated, they are more likely to transfer out of the university entirely. (Leppel, 2002).

For the purposes of this paper, the authors have focused on the level of satisfaction among internal transfer students. We were looking for specific variables that correlate to the rate of internal transfer within the university.

Research Questions:

This study poses three main questions:

1) Is there a correlation between student experiences within the current schools and internal transfer patterns?
    a. More specifically, do student experiences, such as course learning or faculty contact, correlate with the rate of internal transfer?
    b. In addition, are student opinions – regarding the college environment, or estimate of gains, in particular – associated with the number of internal transfers?
2) Do student demographics such as gender, race/ethnicity, SAT scores, or major correlate with the rate of internal transfers?
3) Is there a relationship between transfer patterns and student outcomes?
    a. Specifically, does GPA differ after transferring?
    b. Moreover, what is the rate of persistence of these internal transfers?

In the next section, the methodology will be laid out, discussing the survey instrument and variables used, methodology, and data set. This will be followed by a presentation of the statistical analysis used for determining correlation. Results will then be presented, and finally the authors will discuss possible theories to explain those findings.

Methodology

Participants

The authors researched the responses from the population of matriculated undergraduates at the university whose current class was either freshman or sophomore when completing the College Student Experiences Questionnaire (CSEQ) in Spring 2001. Subsequently, this population was reduced to three study groups.

The authors traced all student respondents' identification numbers from the survey to the institution's student information system to check the number of internal transfers for two categories of transfers. Students transferring prior to/within spring 2001 and those transferring after spring 2001 made up the two study groups, respectively. The authors performed an initial analysis on the whole population of students who were sent the survey instrument to uncover any response bias and ensure that transfers were properly represented in the study groups.

Demographic characteristics were also tracked, such as race/ethnicity, major, current school attended, SAT scores, and grade point average. Cross tabulations were calculated to verify whether any particular characteristic stood out. Since calculated standard deviations were minimal, the authors then pulled a random sample control group.

- Group A: Out of the total population, only those students who responded to the CSEQ and transferred prior to and within spring 2001. A total of 89 students comprised this group.
- Group B: Out of the total population, only those students who responded to the CSEQ and transferred after spring 2001. A total of 67 students comprised this group.
- Group C: A representative sample of students from the total population of CSEQ respondents. A total of 90 students comprised this group.

Instrument

The authors studied a set of questions from the College Student Experiences Questionnaire (CSEQ) to predict a correlation between internal transfer patterns and student experiences.

The survey was administered electronically to all freshmen and sophomores in the spring 2001. There are eight undergraduate schools within the university, and all encompass different school cultures. The authors chose questions on the survey that related closely to experiences found only in the students' particular school of study, rather than the university as a whole. The following items and corresponding responses were selected and examined from the CSEQ:

1. **Course Learning**
   a. Completed the assigned readings for class.
   b. Took detailed notes during class.
   c. Contributed to class discussions.
   d. Developed a role-play, case study, or simulation for a class.
   e. Tried to see how different facts and ideas fit together.
   f. Summarized major points and information from your class notes or readings.
   g. Worked on a class assignment, project, or presentation with other students.

h. Applied material learned in a class to other areas (your job or internship, other courses, relationships with friends, family, co-workers, etc.).

i. Used information or experience from other areas of your life (job, internship, interactions with others) in class discussions or assignments.

j. Tried to explain material from a course to someone else (another student, friend, co-worker, family member).

k. Worked on a paper or project where you had to integrate ideas from various sources.

2. Experiences with Faculty

a. Asked your instructor for information related to a course you were taking (grades, make-up work, assignments, etc.).

b. Discussed your academic program or course selection with a faculty member.

c. Discussed ideas for a term paper or other class project with a faculty member.

d. Discussed your career plans and ambitions with a faculty member.

e. Worked harder as a result of feedback from an instructor.

f. Socialized with a faculty member outside of class (had a snack or soft drink, etc.).

g. Participated with other students in a discussion with one or more faculty members outside of class.

h. Asked your instructor for comments and criticisms about your academic performance.

i. Worked harder than you thought you could to meet an instructor's expectations and standards.

j. Worked with a faculty member on a research project.

3. Opinions About Your College or University

a. How well do you like college?

b. If you could start over again, would you go to the same university you are now attending?

4. Opinions About The College Environment

a. Emphasis on developing academic, scholarly, and intellectual qualities.

b. Emphasis on developing aesthetic, expressive, and creative qualities.

c. Emphasis on developing critical, evaluative, and analytical qualities.

d. Emphasis on developing vocational and occupational competence.

e. Emphasis on the personal relevance and practical value of your courses.

f. Relationships with administrative personnel and offices.

g. Relationships with faculty members.

5. Opinions About Estimate of Gains

a. Acquiring knowledge and skills applicable to a specific job or type of work (vocational preparation).
b. Acquiring background and specialization for further education in a professional, scientific, or scholarly field.
c. Gaining a broad general education about different fields of knowledge.
d. Gaining a range of information that may be relevant to a career.

Scales for responses:

- Course Learning, Experiences with Faculty, and Opinion About Estimate of Gains: 1) Never 2) Occasionally 3) Often 4) Very Often
- Opinions About College Environment: scale of 1 through 7 1= Weak Emphasis, 7=Strong Emphasis)

Procedure

Satisfaction levels were obtained by calculating frequencies of responses. Using the Pearson Chi Square Statistic, correlation rates of student responses to rate of internal transfer were attained. Gender, race/ethnicity, SAT scores, and major demographics were studied for correlation against rate of internal transfer as well.

In tandem, grade point average and persistence rates after one year of transfer were examined to confirm similarities or differences among the control and study groups.

Results

The results indicate that students who internally transferred were indeed similar to those individuals who remained in their original college. The only items on the CSEQ showing any significant correlations using the Pearson Chi Square statistic were:

1. Tried to explain material from a course to someone else, $\chi^2_{df=6} = 16.505, p = .011$: The data demonstrate that Group B (Students who transferred subsequent to taking the CSEQ) was more likely to occasionally explain course material to others as compared to Group A (Students who transferred prior to taking the CSEQ) and Group C (Control group) (41.9% versus 33.9% and 24.2%, respectively). In addition, students in Group B were less likely to explain coursework very often to others, than students in the other groups (22.0% versus 37.8% in Group A and 40.2% in Group C).
2. Acquiring knowledge and skills applicable to a specific job or type of work, $\chi^2_{df=6} = 14.000, p = .030$: The statistics illustrate that Group B were most likely to hold the opinion that very little of its college experience increased the vocational knowledge and skills as compared to Group A and Group C (40.7% versus 29.6% and 29.6%, respectively). As well, students in Group B were less likely to think that they

gained very much vocational preparation compared to the other two groups (22.4% versus 36.2% and 41.4%, respectively).

The Pearson Chi Square Statistic was also run against other student demographics, such as gender, major of study, and SAT scores. The only characteristic that demonstrated a correlation to the rate of transfer was race/ethnicity ($\chi^2_{df=10} = 20.005, p = .029$). Students in Group A were more likely to be Asian American compared to Group B and Group C (45.5% versus 34.1% and 20.5%, respectively). As well, students in Group B were less likely to be non-resident aliens compared to the other two groups (15.0% versus 50.0% and 35.0%, respectively). Other ethnic groups had too few participants to be considered statistically valid in studying correlation.

In addition, student outcomes, such as GPA and graduation rates were studied. The authors discovered that the GPA did not differ after transferring (Pearson Chi Square Statistic: $\chi^2_{df=2} = .112, p = .945$). The students in all three groups maintained similar GPAs overall.

The authors analyzed the five-year historical persistence rates for both internal transfers and students who remained in their original college. Graduation rates from both groups remained relatively consistent across the time period of fall 1992 through fall 1996. Four-year graduation rates for the original group were higher (average four-year rate of 82% for original group versus 76% of internal transfers); however, the internal transfer group had higher six-year persistence rates (average six-year rate of 93% for internal transfers versus 90% of original group).

The authors studied the percentage of transfers within each cohort and discovered that these rates also remained stable across the time frame of fall 1992 through fall 1996. The percentage of those students who never transferred was 88% in fall 1992 to 90% in fall 1996 (average four-year rate of 89%). All in all, the correlation statistics displayed that students across all three groups within the study possess more similarities in experiences, opinions, and demographics than differences. All other CSEQ responses and student demographics produced statistically insignificant correlation rates.

Discussion

In analyzing the correlation rates of responses from the CSEQ to transfer rate, it is interesting to note that Group B was more likely only to occasionally explain course material to others; conversely, Group B was less likely than the other two groups to explain coursework very often to others. As well, Group B was more likely to hold the opinion that very little of its college experience increased vocational knowledge and skills. Students in Group B were also less likely to think that they gained very much vocational preparation.

The authors theorize that those students in Group B are more inclined to be dissatisfied with their course learning, since these are students who transferred *after* responding to the CSEQ. If these students are disappointed with their courses, then they are less likely to be engaged, and therefore less likely to explain course material to others. As well, perhaps these students felt dissatisfaction because they believed the coursework was not relevant to their career paths. In contrast, it might be possible the courses taken were too general to be considered good vocational preparation, since a significant portion of coursework taken by freshmen and sophomores can be considered broad in nature. However, the responses to these two CSEQ questions do display a significant correlation to rate of transfer and should be studied further.

Race/ethnicity seems to play a role in internal transfer according to the findings. Students who transferred prior to responding to the CSEQ were more likely to be Asian American compared to other ethnic groups. A few explanations might point to expectation, satisfaction, and culture of the Asian American community in general. The authors theorize that expectations for the college experience may be different for Asian Americans than other ethnic groups. Asian Americans may expect more interaction with faculty, superior courses, and excellent college facilities than other peers. In tandem, Asian Americans may also feel less satisfaction with faculty and coursework due to these possible higher expectations. Both expectation and satisfaction levels may stem from the Asian American culture. In general, Asian Americans value education and academic achievement and enforce strict rules for reaching this type of success. As a result, they may anticipate more from education, and thus, require higher standards in the college experience overall. Increased expectations may also provide a good indication that Asian Americans are more likely to internally transfer early in their college careers once they determine the initial school is not a good fit.

Students who internally transferred after responding to the CSEQ were less likely to be non-resident aliens. One possible explanation is that this group has a limited amount of time to study in the United States. Therefore, it is less likely that non-resident aliens transfer after their sophomore year, as doing so will generally prolong their time for study, i.e., taking more courses to fulfill credit requirements for graduation. Overall, race/ethnicity and international citizenship may serve as important factors in predicting the rate of internal transfer. Further analysis is needed.

The authors did originally speculate that cumulative GPAs would be lower for internal transfers due to less satisfaction with the courses and faculty. However, no correlation existed between cumulative GPA and rate of transfer. It is theorized that once transferred, those students were more satisfied with their courses and thus, more able to maintain the same cumulative GPA level as those students who remained in their original college. Another factor may be the relatively high GPA requirements for internal transfers as compared to the GPAs required to maintain satisfactory academic study. Most likely, though, it may be that GPA does not play a role in internal transfers. Change of interest may dictate internal transfer, rather than a lower GPA.

The authors assume that this change of interest and thus, better fit for internal transfers may explain the higher six-year graduation rates compared to those students remaining in their original college. Internal transfers may be more engaged, motivated, and satisfied, and therefore have higher persistence rates. Perhaps transfers also believe that they were able to control the direction of their college career, and have a greater stake in their educational outcome. Transfers may feel it is important to see their decision through to a conclusion. This may also explain the reason graduation rates are higher for this group. In addition, the internal transfer process is quite difficult at this institution with barriers in place that discourage internal transferring. Therefore, students who do not overcome those hurdles remain in the original college or may not persist at all. This may also contribute to the lower six-year graduation rate for the original group. Since the historical rates for internal transfers and graduation remain relatively consistent, total undergraduate enrollment may be better predicted and managed in the future.

Overall, what are the implications of the research findings? The authors note that the university is not in a position to forecast accurate data for enrollment management practices at the individual school level due to the fact that the transferring and non-transferring students are similar. On the other hand, the university might improve enrollment projections on a macro level by incorporating the stronger correlating factors such as ethnicity/citizenship and stable historical trends. In addition, time to degree is probably minimally affected and might be more accurately predicted due to consistent internal transfer and graduation rates. Further study of CSEQ responses may also uncover other important factors for projection data.

Enrollment management of internal transfers will continue to remain problematic on an individual school basis since there are no outlying differences among students who internally transfer and those who remain in their original school. In contrast, this is more likely to be tempered by the decentralized internal transfer process itself. Thus, financial aid models and resources are likely to be minimally affected because students are more likely to stay within their original college and obtain the same financial aid packages. As well, due to the similarities among the three groups, the research indicates most of these internal transfers remain a good fit for the university as a whole. This finding reassures the university that most students will stay and not leave the institution prior to obtaining a degree. Therefore, the authors speculate that the rate of internal transfers will not contribute to an increase in university-wide attrition rates.

Why do the three groups of students seem more similar than different? The two groups of transfer students are generally satisfied with their college experience, both with faculty and course learning. Originally, it was theorized that students internally transferring subsequent to responding to the CSEQ would have been less satisfied. With a couple of exceptions, the findings indicate that both groups of transfers were content with their college experience compared to the control group. The authors now speculate that most students indeed are partial to the university. Perhaps they value the status that the university conveys and desire a degree from the university. Or, possibly, the students believe they are receiving a good education, and instead of externally transferring, decide to transfer within the university to match their change of interests. Plus, the possibility of embarking on a different course of education without moving to another institution is likely a boost to overall institutional retention.

Conceivably, internal transfers may already be socially integrated. These students have already invested time and effort into social relationships. To support this assumption, within this particular university, friendships do not have to remain within the individual student college; students are integrated throughout the university by means of housing, activities, etc. If students believe they are vested in these relationships, they may transfer within instead of outside the university. Social integration should be further studied to better understand the factors influencing internal transfers.

As mentioned before, the level of coursework may be too general in nature to be considered good vocational preparation. In contrast, the courses may be general enough that students believe their learning is applicable to their occupation. Perhaps this mainly identifies a reason as to why the groups are not that different.

After studying the results, the question remains: Why do students internally transfer? It is speculated they do simply because of a change in interest and career. Approximately one in every 10 students internally transfers; this rate may correlate with the number of students who simply change their minds. Since only correlations were examined, causes for internal transfer need to be studied further.

Conclusion

The study was undertaken to better understand enrollment patterns at a Research I University. The challenges include producing more reliable models for enrollment, financial aid, time to degree, and attrition rates. Internal transfers have never been easily predictable. The authors theorized this group may have added to the challenges in modeling and therefore sought to uncover any factors that would contribute to modeling more accurately internal transfers.

The authors examined the group of freshmen and sophomores who responded to the CSEQ to find links between internal transferring and level of satisfaction with college experience. Three groups of students were studied: those who transferred prior to responding to the CSEQ; those who transferred after responding; and a control group consisting of a representative sample of respondents who did not transfer.

We examined student responses to an internal survey, student outcomes, and historical internal transfer patterns. Correlation rates to transfer rates were compared using student responses, regarding faculty interaction, course learning, and overall satisfaction level; student demographics, such as gender, race/ethnicity, SAT score, and major of study; and student outcomes, such as GPA and persistence rates.

The study revealed only two items on the administered student survey that correlated significantly to the rate of internal transfer. Students who transferred subsequent to responding to the survey were more likely only to occasionally explain course material to others compared to other groups. In addition, this same student group was more likely to think that very little of the college experience increased vocational knowledge and skills than the other two groups. Reasons for these differences could be further studied.

The research also uncovered that race/ethnicity correlated significantly with the rate of internal transfer. Asian Americans were more likely than other ethnic groups to transfer in the first two years of their academic career. In contrast, non-resident aliens were less likely to transfer after their sophomore year. Further study could be undertaken to examine the different ethnic populations and the reasons affecting satisfaction level, so that enrollment models might accurately reflect current trends.

As well, the study revealed that the percentage of internal transfers continued at a stable rate of approximately 11% over a five-year period and that five-year historical graduation rates remained steady for internal transfers over the same period. Students who internally transfer had a slightly higher six-year persistence rates than the original group (93% compared to 90%). Due to the relative stability, macro-level models can be developed to forecast enrollment, financial aid, time to degree, and attrition rates using certain data for the group of internal transfers.

Overall, the research indicated that all three groups of students possessed more similarities than differences in demographics, experiences, and opinions. The study revealed that most students are indeed a good fit for the university. Thus, the group of internal transfers can be accurately projected for enrollment, financial aid, time to degree, and attrition rate models at the macro level. In contrast, the individual schools will still find it problematic to forecast this student population for their own internal purposes.

Works Cited and References

Alvord, Catherine J. (2001). Undergraduate Enrollment Trends – Fall 2001. *http://www.ipr.cornell.edu/Reports/report_menu.htm.*

Leppel, Karen (2002). Similarities and Differences in the College Persistence of Men and Women. *The Review of Higher Education, Summer 2002, 25*(4), 433-450.

AN ANALYSIS OF THE RETENTION OF FIRST TIME FULL TIME FRESHMEN AT A PUBLIC URBAN UNIVERSITY

Kevin B. Murphy
Research Analyst
The Office of Institutional Research and Policy Studies
University of Massachusetts Boston

Introduction

The one year retention and persistence to a degree of first time full time college students has been studied at length. However, one problem with these retention studies is that they tend to deal with retention and persistence as a "one size fits all" proposition. While public urban universities have been organizing themselves into a number of consortia to both identify the differences in their missions versus those of their public flagship university counterparts, and to stress the differences in the populations the public urban universities tend to serve, research into what those differences might mean for both retention and persistence rates has not been as actively pursued.

Our public urban university has recently conducted a study of the one-year retention of fall 2000 first time full time freshmen. Although this group of students made up only about 15% of all newly admitted students in the academic year, they are particularly important because they are the group for which retention and persistence rates are reported. This study focuses on entrance and demographic variables from the university's databases, and on responses to a telephone survey of both retained and non-retained students that was conducted by a professional surveying entity.

The study has identified a number of issues and concerns. As we analyzed the results, we found that a number of factors that are theorized to affect retention in the retention literature either have little impact on this group of public urban university freshmen, or in some cases, have an effect that is the opposite of that which is theorized. Among these factors are race/ethnicity, Verbal SAT scores, and a number of variables that reflect institutional commitment. At least part of these differences may well be related to our status as a public urban university where race/ethnicity, citizenship and language status, Verbal SAT scores, and institutional commitment may all be related in a number of ways. Further, for freshman entrants, all of these elements are related to participation in a pre-matriculation summer program offered on campus for a selected group of students who lack the traditional credentials for admission to a four year college, but who show promise of benefiting from a university education. The program is called Directions for Student Potential (DSP) and it provides intensive workshops in reading, writing, English language, mathematics, and study skills. Those students who successfully complete the program are admitted to the university's College of Arts and Sciences (CAS) in the fall. Students who had completed the program were retained at a rate that is significantly higher than that of non-participants. The difference of almost 12% was significant at the 99% level.

The study was developed because of ongoing concerns regarding the university's retention rates. While public urban universities in general tend to have lower retention rates than their flagship counterparts, the university's fall 2000 cohort retention rate ranked ninth in a set of twelve when compared to eleven other public urban universities that the state university system office was using for peer comparisons at that time.

Data and Methodology

The data used for this analysis are drawn from two main sources. The first is official university admissions and enrollment data were accessed for all 595 of the fall 2000 first-time full-time freshmen; the second is a telephone survey of both returning and non-returning students in that cohort was conducted by the Center for Survey Research (CSR) at the University of Massachusetts Boston. Additional historical data have been drawn from the university's files, and data from the 2002 administration of the National Survey of Student Engagement have also been used to inform the discussion of retention at the university.

By the beginning of October 2001, 411 of the 595 first-time full-time freshmen from fall 2000 had returned. We decided to survey all of the 184 non-returning students and a randomly drawn sample of 200 of the 411 returning students. After pre-testing the instrument with both groups, CSR prepared the final survey instrument. CSR received responses from 75 (41%)of the 184 non-returning students and from 135 (62.5%)of the 200 returning students in the random sample of returning students. Overall, 35% of the total cohort responded to the survey.

Compared to their proportion in the cohort, men were slightly overrepresented among respondents to the survey, probably because they were overrepresented in the non-returning group. The racial/ethnic mix of the respondents was roughly equivalent to that of the overall group, with Asian/Pacific Islanders and international students slightly underrepresented, and White non-Hispanics slightly overrepresented.

This analysis primarily deals with descriptive statistics concerning a limited number of variables. For a number of these variables, comparison of means tests will be reported, and in one case, a simple bivariate logistic regression model was used and reported. The use of the bivariate logistic regression model was intended to make the strength of the relationship more accessible to the reader. It is not meant to be a final model. More work needs to be done to access and analyze more variables for the specification of a more complete model, and to deal with the high degree of correlation among a number of variables such as race/ethnicity, whether English is spoken at home, participation in the DSP program, Verbal and Math SAT scores, etc. The details of the correlation of several key variables are presented in Table 1.

Table 1: Correlation of Key Variables

	satmath	satverb	us_citizen	dsp	eng_at_home
satmath	1.0000				
satverb	**0.4377**	1.0000			
us_cit	0.0114	**0.4969**	1.0000		
dsp	***-0.4302***	***-0.5326***	-0.1296	1.0000	
eng_at_home	0.1405	**0.4405**	**0.4671**	-0.1729	1.0000
asianpac	0.0878	-0.2229	-0.2072	0.1249	-0.3518
blacknh	-0.2499	-0.1626	0.0282	0.1962	-0.0083
hispanic	-0.2530	-0.1312	0.0061	0.2151	-0.2396
whitenh	0.1711	**0.4061**	0.3878	-0.2130	**0.5542**

Results

Gender and Race/Ethnicity

Men made up about 45.2% of the fall 2000 entering cohort of first time full time freshmen. This is not markedly higher than the 42.6% of the matriculated undergraduate headcount that men represent. Of the 269 men who entered in fall 2000, 173 (64.3%) returned in the fall of 2001. For women, the retention rate was much better; 238 of the 326 women (73%) who entered in fall 2000 returned to the university in fall 2001. A comparison of means test on the difference of about 8.7% returned a T-statistic with an absolute value of about 2.3, which is statistically significant at the .05% level.

There was also variation in the retention rate by racial/ethnic group. Among the larger groups, the Asian/Pacific Islanders returned at a rate that was significantly higher than either the overall group mean or the non-Asian/Pacific Islander group. The White non-Hispanic group returned at rate that was significantly lower than both the overall group mean and the mean for those not within the White non-Hispanic group. While there is some variation for the other groups, it is not statistically significant.

Most of the strength of the gender association can be found in just two of the racial/ethnic groups. For the Asian/Pacific Islanders and the White non-Hispanics, women were significantly more likely to be retained than were men. There are no statistically significant differences in the other groups. Aside from the Native Americans, the lowest retention rate by race/ethnicity and gender was for White non-Hispanic males, only about 56% of whom returned for fall 2001. Details of retention rates by race/ethnicity and gender are presented in Table 2.

Table 2: Comparison of Retention Rates by Race/Ethnicity and Gender

Racial/Ethnic Group	**% Retained Total Cohort**	**Women**	**Men**	**Difference**
Asian P/I	83%	91%	75%	16%
Black N/H	76%	78%	71%	7%
Cape Verde	100%	100%	NA	NA
International	73%	67%	81%	**-14%**
Table 2: cont.	% Retained Total Cohort	Women	Men	Difference
Hispanic	64%	60%	71%	**-11%**
Unk/Refused	72%	78%	66%	12%
White N/H	62%	67%	56%	12%
Total	**69%**	**73%**	**64%**	**9%**

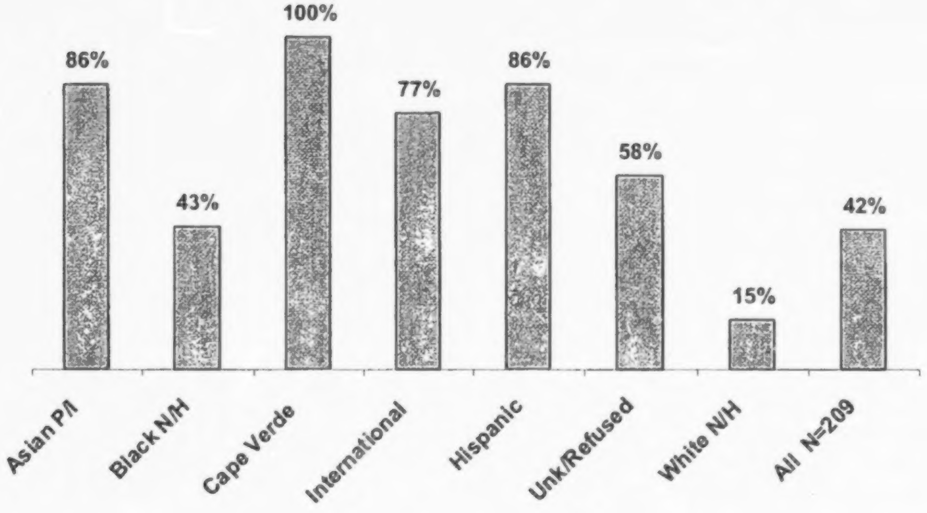
While we report these racial/ethnic categories as if they are homogeneous, we found that for the most part, they are not. There is much cultural diversity within the groups. Except for the Cape Verdean and Native American groups, university records showed that there was a tremendous amount of variation within the groups by citizenship status. Citizens of countries other than the United States made up more than 50% of the Asian/Pacific Islanders, more than 25% of Black non-Hispanics, more than 28% of Hispanics, more than 10% of the White non-Hispanics, and almost 30% of those who had refused to supply racial/ethnic information. This does not overlap those who are here as international students. Overall, the 595 students in the cohort claimed citizenship in more than 60 countries. Non-U.S. citizens made up about 30% of the cohort as a whole. These non-U.S. citizens were retained at a 77% rate, while U.S. citizens were retained at a 60% rate. The difference is statistically significant at the .05% level.

Further evidence of the cultural diversity within the racial/ethnic categories required for federal and state reporting came from the language questions asked of all returning and non-returning students. The students were asked whether they usually spoke a language other than English at home or with family. Those who responded that they did speak a language other than English at home or with family were then asked which language. Overall, about 42% of respondents reported speaking a language other than English at home. The percentages of each racial/ethnic group who reported using a language other than English at home are presented in Chart 1.

The racial/ethnic category of "Unk/Refused" indicates a refusal by the student to supply the information. At the university, it has often been assumed that this category is primarily made up of White non-Hispanic students. However, both the nations of citizenship and the languages spoken at home by this group challenge that assumption.

This diversity within the standard reporting categories is important because it affects assumptions we may make about culture and native language status based simply upon standard racial/ethnic categories. Programs that are designed to increase retention should take this cultural and linguistic diversity into account. The degree of cultural and linguistic diversity was surprising to some people at the university. However, the percentage of respondents who reported speaking a language other than English at home is consistent with percentages suggested by responses to an Urban Consortium question in the 2000 administration of the National Survey of Student Engagement, several small surveys of First Year Seminar participants, and most recently, a survey of students who had applied for graduation in August 2002.

Chart 1: Percent of Survey Respondents Who Speak a Language Other than English at Home or with Family
By Race/Ethnicity

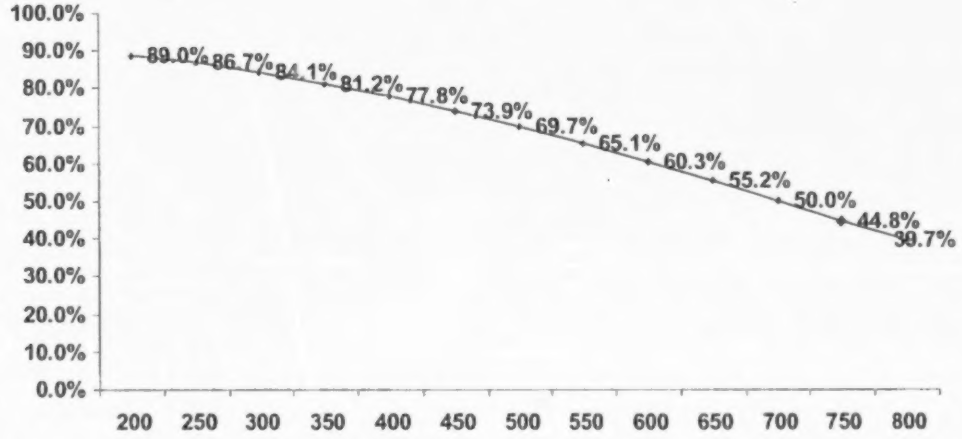


SAT Scores

Among all 595 first time full time freshmen, including those who entered through the DSP program, we have SAT scores for 491. Several tests of significance were run to examine the relationship between the Math SAT scores and retention. No statistically significant relationship was found. However, when the same tests were run for the Verbal SAT scores, a very strong relationship was found. A simple significance of difference test was run on Verbal SAT score by retention status. The results showed that retained students had a mean score of about 463, while non-retained students had a mean of about 510. The difference of 47 points was significant at the highest level.

A simple bivariate logit regression was then run, with the dummy variable for retention as the dependent variable and the Verbal SAT score as the independent variable. The results were statistically significant at the highest level. The predicted values obtained from the regression results show that, on average, for every 50 points that the Verbal Sat score rises, the probability that the student will be retained drops by an average of 4 percentage points, and that the rate of attrition accelerates as the Verbal SAT rises. Details are presented in Chart 2.

Chart 2: Predicted Retention Rate
By Verbal SAT Score



Part of this relationship may be explained by race/ethnicity. There are marked differences in Verbal SAT score by race/ethnicity, and as already noted, differences in retention rates by race/ethnicity. Details are presented in Table 3.

Table 3: Significance of Verbal SAT Differences by Racial/Ethnic Group

Race/Ethnicity	Mean Verbal Score	Difference from Overall Mean	Statistical Significance Level	Group Comparison Mean	Difference	Statistical Significance Level
Asian P/I	426	-51	>99%	488	-62	>99%
Black N/H	423	-54	>99%	483	-60	>99%
Cape Verde	380	-97	>95%	478	-98	>95%
Foreign	393	-84	>99%	480	-88	>99%
Hispanic	425	-52	>99%	481	-88	>99%
Unk/Refused	480	3	None	476	4	None
White N/H	523	46	>99%	435	89	>99%

Another contributing factor may be language status. Those who spoke a language other than English at home had Verbal SAT scores that were, on average, about 93 points lower than those who spoke English at home. However, their retention rate for these students, while about 9.5% higher than the rates of students who spoke English at home, was not significantly different at even the .05% level. The lack of statistical significance may be because of the small numbers available for analysis. When U.S. citizenship (obtained from university records) is used as a proxy for language so that we could analyze the entire cohort, we found that U.S. citizens had mean SAT scores that were over 120 points higher than the non-citizens, and that the non-citizens were over 18% more likely to be retained. Both of these differences were significant at the highest level.

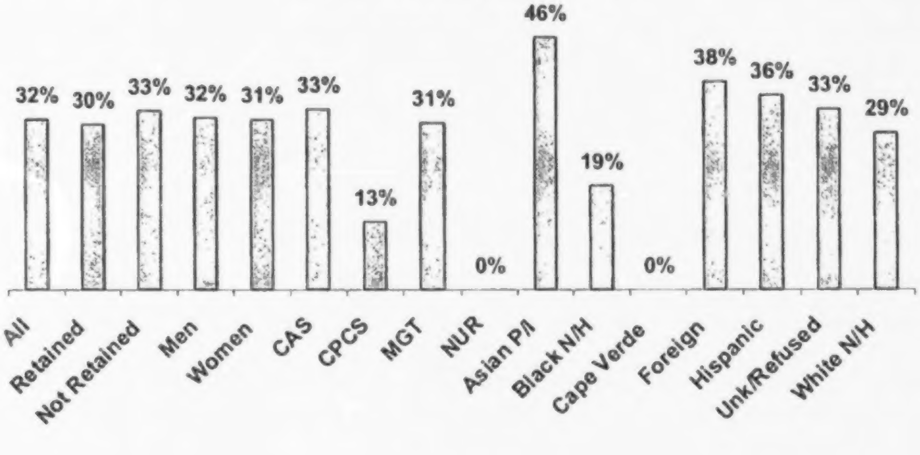
Another possible confounding factor that was examined was participation in the Directions for Student Potential program. DSP students had significantly lower Verbal and Mathematics SAT scores and significantly higher retention rates.

Regardless of the confounding factors, this negative relationship between Verbal SAT scores and retention was unexpected.

Institutional Commitment

All of the respondents were asked if, when they first enrolled at the university, they intended to obtain a degree here or not. About 32% of the respondents said that they were not planning to get a degree here. There was remarkably little variation in the mean of the responses to this question across a number of categories. Details of the survey responses about plans to complete a degree at the university when first enrolling are presented in Chart 3.

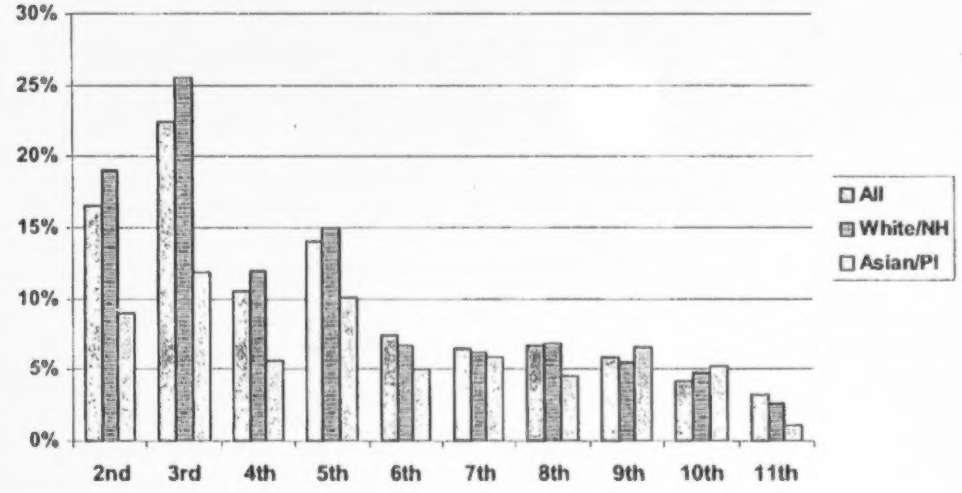
Chart 3: Percent of Survey Respondents Who Were Not Planning to Complete a Degree at the University When They First Enrolled



The only difference of statistical significance is for Asian/Pacific Islanders, about 46% of whom were not planning to get a degree from the university when they entered, compared to about 29% of the non-Asian/PacificIslanders. Given that the Asian/Pacific Island group is the one most likely to be retained for the first year, this pattern immediately raised fears that their attrition will be rapidly accelerated as they approach upper division standing. In a previous internal university study, the author noted that Asian/Pacific Island students had a higher one year retention rate than students in other racial/ethnic categories, but that their persistence rates began to sharply decline after the fourth semester. She hypothesized that they drop out for different reasons than do other students and called for more research on the subject. She also cited anecdotal evidence from faculty and staff that this represented transfer out behavior that was planned prior to first entry at the university as part of a plan to transfer to and receive a degree from a more prestigious institution. There are competing hypotheses.

However, the idea that the Asian/Pacific Islander students have good initial retention only to leave the university after the fourth semester seems to be a widely held university myth. Chart 4 reports the mean percentage loss of students by semester for fall freshmen entrants from 1990 to 2000. This group loses a lower percentage of its members than do the white non-Hispanics or the overall group every semester through the ninth semester after first enrollment. Not until the ninth semester do we observe even a slightly higher one semester attrition rate for Asian/Pacific Islander students (7% for Asian/Pacific Islander vs. 6% for White non-Hispanic and the overall group). At the ninth semester, on average 54% of the Asian/Pacific Islander students are still enrolled or have graduated vs. 38% of the group as a whole and 35% of white non-Hispanics.

Chart 4: Mean Percentage of Remaining Students Lost From The Previous Semester for 1990-2000 Fall Entering Freshmen By Race/Ethnicity and Semester from First Enrollment



A general incoming plan not to complete a degree at the university does not seem to be affecting retention yet, because the percentage of students reporting that they were not planning to get a degree was essentially the same for both the retained and non-retained groups.

In addition to the direct question about whether the student was planning to get a degree at this university when first enrolling, the survey asked all of the students about eleven specific possible reasons they may have had when they made the decision to attend the university. Each reason was offered as a separate question with a "yes" or a "no" response requested. They were to select each reason that applied to them. The wording of the introduction to the section was "I'd like you to think back about why you might have decided to attend UMass Boston. For each of the following, please tell me if it was one of your reasons for coming". The reasons offered were:

1) To Obtain a Degree
2) Cost
3) Location
4) Develop Work Skills
5) Academic Reputation
6) Specific Course of Study
7) Student Diversity
8) Quality of Instruction
9) Get Credits for Transfer
10) Friends or Family Attend UMB
11) Didn't Get into 1st Choice

Nine of the 11 reasons are generally positive with regards to retention in that they provide a reason for the student to come to and to stay at this university. However, two of the reasons are decidedly negative in terms of potential retention. Both "Get Credits for Transfer" and "Didn't get into 1st Choice" demonstrate a lack of commitment to attending the university. Institutional commitment has been found to be a key variable in a number of previous retention studies (Bean, 1979; Brower, 1992; Cabrera et al, 1993). Indeed, it is a key factor in both Tinto's model of student integration, and in Bean's model of student attrition (Bean, 1979; Tinto, 1975 & 1993).

The two "negative" reasons of "Get Credits for Transfer" and "Didn't Get into 1st Choice" have exactly the effect one would expect. Those who expressed them were more likely to leave the university than those students who did not. An answer of "Yes" to either of these questions is evidence of a lack of institutional commitment, but the real impact seems to come in the combination of the two. If the students were rejected by their first choice and came with the plan to get credits for transfer, then we retained them at only a 45% rate. While the difference was not statistically significant, the mean Verbal SAT scores of the students who chose both of these reasons and left was 70 points higher than that of the students who selected both of these negative reasons and returned.

The details of the relationships between selecting these negative reasons and retention are reported in Table 4.

Table 4: Impact of Selecting "Negative" Reasons on Retention Rates

Negative Reason	N Size	Retention Rate	Difference from Overall Mean	Statistical Significance Level
Neither Selected	118	70.3%	5.7%	Not significant
Get Credits for Transfer Only	34	61.8%	-2.8%	Not significant
Not 1st Choice Only	23	60.9%	-3.7%	Not significant
Both Selected	31	45.2%	**-19.4%**	95%

Conclusions

Study of the fall 2000 cohort of first time full time freshmen at the university is ongoing. We are planning to track the same students into the third year with updated information from university databases, and we intend to attempt to conduct several focus groups with retained and non-returning students. Further, we intend to access the same university data on the fall 2001 entering cohort in an attempt to identify whether the patterns that were observed for the fall 2000 cohort and are reported here were an aberration.

It was certainly unexpected that the retention rate for White non-Hispanic students would trail the retention rate of all other students by 14 percentage points, that the relationship between Verbal SAT scores and retention was not only negative, but strongly so, that those who were immigrants and/or spoke a home language other than English would have retention rates higher than citizens and home English speakers, and that the group that was least expecting to graduate from the university when they first enrolled would have the highest one year retention rate.

Public urban universities have been having a conversation for some time about the fact that they differ from their public flagship or private counterparts in the diversity of the students they serve. In the ongoing analysis of retention at this public urban university, identifying important factors that affect retention has been difficult because of the degree of diversity within the more obvious diversity. Susan Choy's recent report for the American Council on Education identified understanding the recent explosion of diversity in college populations as essential to the appreciation of access and attainment in higher education (2002). Public urban universities are in the forefront of dealing with this diversity, and if we are to not only provide access to higher education but the appropriate tools to be able to succeed in higher education, we will need to pay attention and look below the surface diversity to be able to identify the cultural and linguistic diversity within the categories with which we have become comfortable.

References

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education, 12* (2).

Brower, A. (1992). The "second half" of student integration: The effects of life task predominance on student persistence. *The Journal of Higher Education, 63(4),* 441-462.

Cabrera, Alberto F., Nora, Amaury, and Casteñeda, Maria B. 1993. College Persistence: Structural equations modeling test of an integrated model of student retention. *The Journal of Higher Education ,64(2):* 123-139.

Choy, S.P. (2002), *Access & Persistence: Findings from 10 years of longitudinal research on students.* Washington D.C.: American Council on Education.

Tinto, V. (1975). Dropouts from higher education: A theoretical synthesis of the recent literature. *A Review of Educational Research, 45*, 89-125.

Tinto, V. (1993). 2nd Edition. *Leaving college: Rethinking the causes and cures of student attrition.* Chicago: The University of Chicago Press.

THE IMPACT OF CONTACT TYPE ON WEB SURVEY RESPONSE RATES

Stephen R. Porter, Ph.D.
Director of Institutional Research
sporter@wesleyan.edu

Michael E. Whitcomb
Assistant Director of Institutional Research
mwhitcomb@wesleyan.edu

Office of Institutional Research
237 High Street
Wesleyan University
Middletown, CT 06459

Introduction

Web surveys are becoming more and more popular, as reflected in the growing research in a variety of fields on web survey methodology and response rates (e.g. Cho and LaRose 1999; Cobanoglu, Warde, and Moreo 2001; Cook, Heath, and Thompson 2000; Couper 2000; Couper, Traugott, and Lamias 2001; Crawford, Couper, and Lamias 2001; Dillman and Bowker 2001; Dillman et al. 1998; Heerwegh and Loosveldt 2002; Pealer et al. 2001; Shannon et al. 2001; Sills and Song 2002; Tourangeau, Couper, and Steiger 2001). Despite this spate of recent research, it is still not entirely clear if the techniques used to increase response rates in paper and telephone surveys will directly translate to surveys conducted via the Internet. Yet understanding which techniques increase response rates in web surveys is increasingly important. Researchers are faced with decreasing response rates in surveys (Smith 1995; Steeh 1981), as well as increased competition with marketers and spammers on the Internet for the cooperation of respondents (one recent study predicts that by 2006 email users will receive over 1,400 spam messages per year (Tynan 2002)). Without such knowledge, web surveys may become less useful as a tool for survey research.

Previous research on web surveys has focused on areas such as coverage and sampling error, privacy issues, effects of multiple follow-ups, controlling survey access, and survey appearance (see Table 1). Very little research has been conducted on how the process by which members of the sample are contacted affects the probability of response. This is in direct contrast to the rich literature on method of contact in paper and telephone surveys (e.g. Dillman 1991; Dillman 2000). Factors such as the approach technique of the interviewer, the appearance of the survey envelope, the salutation of the survey cover letter, the interviewer script, and the content of the cover letter are all correlated with survey response rates. What we now need is a similar body of research describing how aspects of the initial and follow-up contacts in electronic surveys affect survey response.

Without a deeper understanding of the contact and response process, researchers may use techniques that are time-consuming and ineffective.

Table 1. Summary of Recent Research on Web and Email Surveys

Approach	Topic	Studies
Experimental	Web/email surveys versus fax/telephone/postal surveys	Cobanoglu, Warde, and Moreo 2001; Kiesler and Sproull 1986; Mehta and Sivadas 1995; Pealer et al. 2001; Schaefer and Dillman 1998; Shannon and Bradshaw 2002; Weible and Wallace 1998; Yun and Trumbo 2000; Zhang 2000
	Coverage and sampling error	Bradley 1999; Witte, Amoroso, and Howard 2000
	Embedded vs. attached email surveys	Dommeyer and Moriarty 2000
	Controlling survey access	Crawford, Couper, and Lamias 2001; Heerwegh and Loosveldt 2002
	Issue salience	Manfreda, Batagelj, and Vehovar 2002; Sheehan 2001; Sheehan and Hoy 1999
	Survey appearance	Couper, Traugott, and Lamias 2001; Crawford, Couper, and Lamias 2001; Dillman et al. 1998; Manfreda, Batagelj, and Vehovar 2002; Tourangeau, Couper, and Steiger 2001
	Survey length	Sheehan 2001
	Prenotification contacts	Sheehan 2001; Sheehan and McMillan 1999
	Multiple follow-ups	Crawford, Couper, and Lamias 2001; Kittleson 1997; Mehta and Sivadas 1995; Sheehan 2001
Descriptive	General methodology	Dillman 2000; Vehovar et al. 2002
	Coverage and sampling error	Bradley 1999; Couper 2000; Dillman and Bowker 2001; Sills and Song 2002
	Survey appearance	Dillman and Bowker 2001
	Privacy issues	Cho and LaRose 1999
	Typology for response behaviors	Bosnjak and Tuten 2001

A second reason why understanding how the method of contact affects web survey response is simple economics. Changing the email contact in a web survey is an almost costless technique for increasing response rates. Unlike paper surveys, where techniques such as the use of individual postage stamps versus metered mail, and signing cover letters versus photocopying a signature, can be quite costly, there are generally no differences in cost between different versions of an email contact. As most researchers face limited resources when conducting a survey research project, studying different ways to contact respondents should yield cost-effective ways to increase response rates.

Using a web survey of over 11,000 high school students, we investigated the impact of characteristics of the email contact on response rates, varying such attributes as personalization of salutation, email address, job title and office of sender, statements of deadlines, and statements of selectivity. Our results indicate that some of the tactics used to increase response rates in paper surveys may not directly translate to the electronic realm.

The paper is divided into four parts. After reviewing the relevant literature and discussing how techniques used in paper surveys can be used in web surveys, we describe our survey and experimental design. Next, we describe the results, and discuss our findings and their implications.

Literature review

In their review of why someone decides to participate in a survey, Groves, Cialdini and Cooper (1992), discuss research in social psychology that sheds light on the decision to participate. Besides considering the direct costs and benefits of participation, they argue that people often use a heuristic approach when deciding whether to participate:

> …the survey request situation most often favors a heuristic approach because the potential respondent typically does not have a large personal interest in survey participation and, consequently, is not inclined to devote large amounts of time or cognitive energy to the decision of whether to participate (p. 480).

They describe several areas where people use these heuristic devices, three of which we investigate in this study: reciprocity, authority, and scarcity.

First, under the norm of reciprocity, people feel obligated to respond to gifts or other positive actions from another person (Groves, Cialdini, and Couper 1992). In terms of survey research, this norm explains the success of small pre-paid monetary incentives in garnering large response rates (e.g. Church 1993; Fox, Crask, and Kim 1988; Warriner et al. 1996; Willimack et al. 1995; Yammarino, Skinner, and Childers 1991). A similar rationale probably underlies the common finding that personalization of a survey, such as the use of postage stamps, addresses printed on envelopes rather than the use of mailing labels, actual signatures, and other design techniques to distinguishing the survey from advertising material, generally results in a higher response rate (Dillman 1991; Fox, Crask, and Kim 1988). If survey participants believe the surveyor has spent a lot of time and resources in constructing the survey, they may be hesitant to throw it out and instead decide to participate. Aspects of the contact that indicate investment of resources on the part of the surveyor should invoke the norm of reciprocity, in turn increasing the probability of response.

Second, Groves et al. cite research that indicates people will be more willing to comply with a request when that request comes from a legitimate authority. Government surveys, for example, often have higher response rates than surveys by other organizations (Goyder 1987). Thus interviewers who invoke a survey sponsor who yields legitimate authority in the area of the survey should be more likely to achieve survey participation (Groves, Cialdini, and Couper 1992). A possible corollary is that interviewers with higher-level positions should also have higher success rates for survey participation, although this could in part be due to reciprocal feelings raised by having an important person spending time administering the survey. Empirical evidence on the impact of survey sponsorship is mixed, with some scholars finding positive effects (Fox, Crask, and Kim 1988; Goyder 1987; Heberlein and Baumgartner 1978), and others no effect (Presser, Blair, and Triplett 1992; Yammarino, Skinner, and Childers 1991).

A third heuristic device affecting response rates is the perception of scarcity. In general, people have a tendency to view scarce opportunities as more valuable than more common opportunities. In terms of survey research, statements suggesting that only a few people have been selected to participate should elicit higher response rates (Groves, Cialdini, and Couper 1992). Similarly, giving respondents a deadline should also increase participation, as the time frame for participation is limited. The literature is mixed as to whether deadlines increase response rates in mail questionnaires (Fox, Crask, and Kim 1988; Roberts, McCrory, and Forthofer 1978; Yammarino, Skinner, and Childers 1991; Yu and Cooper 1983).

While there are many studies of contact approach in traditional paper and telephone surveys, experimental research on the impact of contact method in web surveys is limited. The vast majority of research either compares email and web surveys with traditional survey methods or examines various aspects of survey design (see Table 1). Some research has been conducted on number of email contacts (Crawford, Couper, and Lamias 2001; Kittleson 1997; Mehta and Sivadas 1995; Sheehan 2001; Sheehan and McMillan 1999).

In our review of the literature we did not discover any studies examining the effect of personalization, indicators of authority, statements of scarcity, or statements of survey deadlines in electronic surveys, with the exception of the study by Cook, Heath and Thompson (2000). Their meta-analysis of 68 web surveys indicates personalized contacts have a positive impact on response rates.

Web surveys conducted by other researchers indicate a belief that personalized contacts and statements of survey deadlines matter. Although they did not test the impact of personalization per se, some researchers have personalized their email contacts in their studies (Heerwegh and Loosveldt 2002; Mehta and Sivadas 1995; Schaefer and Dillman 1998). Others have included statements indicating a deadline when the web survey would be shut down (Heerwegh and Loosveldt 2002).

In sum, research on survey response in non-electronic surveys indicates that efforts at personalization, survey sponsorship by legitimate authorities, and statements of scarcity may positively affect response rates. Research on whether these methods might affect response rates in electronic surveys, however is quite limited.

Data and Methodology

Data

The study is based on a web survey of high school students who had contacted a selective liberal arts college for information, but did not apply for admission. The survey asked over fifty questions about perceptions of and reasons for not applying to the college. The salience of this survey is low, as evidenced by the 15% response rate in a similar version of the survey administered one year previously. After 189 incorrect email addresses were removed through a software program that checks the validity of each email address on the email server, the remaining sample size was 12,433. The sample sizes for each experimental group are given in Table 2.

All students in the sample were sent an initial email, and nonrespondents were sent up to two follow-up emails. Each email contained a unique URL that automatically logged the student into the survey. Unlike some other researchers (e.g. Mehta and Sivadas 1995; Sheehan and Hoy 1999), we did not receive any accusations of spam for our unsolicited contacts, most likely due to the prior communications by the Office of Admission with the students. The final response rate for the survey was 14.8%.

Methodology

In order to test the effects of personalization, authority and scarcity in email contacts, we developed two experiments. Both experiments test the effects of various components of the email request on survey participation. With web surveys, however, it is possible to go beyond the traditional dichotomy of unit responders and non-responders (Bosnjak and Tuten 2001; Vehovar et al. 2002). We look at both the click-through rate and the response rate, where the click-through rate is defined as the percentage of respondents viewing the first page of the survey, but not submitting any results (as determined by the log files from the server).

Experiment I

The first experiment was 2x2x2x2 design that tested the impact of personalization and authority of the survey administrator. The email contact was varied in four areas:

- The email salutation was either personal (e.g., Dear Jane) or impersonal (Dear Student).

Table 2. Sample Sizes

Experiment I		Authority of requesting office: low		Authority of requesting office: high	
		Source address: office	Source address: person	Source address: office	Source address: person
Authority of signatory: low	Salutation: impersonal	536	544	536	534
	Salutation: personal	541	582	540	546
Authority of signatory: high	Salutation: impersonal	537	549	544	558
	Salutation: personal	519	533[a]	527	547

Experiment II		Selectivity statememt	
		Yes	No
Deadline	No deadline	559	533[a]
	Deadline in email 3	519	543
	Deadline in emails 2 & 3	539	529
	Deadline in emails1,2 & 3	531	540

[a] Same experimental group.

- The email address of the sender was either personal (e.g. jsmith@institution.edu) or impersonal (surveyresearch@institution.edu).
- The authority of the email signatory as reflected by job title was either high (Director), or low (Administrative Assistant).
- The authority of the requesting office was either high (Office of Admission), or low (Office of Institutional Research). Given that the students' past contacts were solely with the Office of Admission, their perception should be that the Office of Admission would have more legitimate authority to request information about why students chose not to apply.

Finally, emails in Experiment I had no mention of a survey deadline or a statement about the selectivity of participation as in Experiment II.

Experiment II

The second experiment tested the impact of perceived scarcity of survey participation on survey responses, using a 2x4 design. Specifically, we varied the inclusion of a selectivity statement and a participation deadline in our email contacts with survey participants. To indicate selectivity of participation, a statement about being part of a select group of students asked to take part in other survey ("You are one of a small group of students who have been randomly selected to provide feedback about our institution.") was either included or excluded from all email contacts requesting survey participation.

For the deadline factor, a statement of the last possible day to participate in the survey ("The website will be closed at midnight on Friday, February 22, 2002.") was either included or excluded from email contacts. We used four different groups to test the impact of the deadline statement. Students in the no deadline condition were never informed of the survey closing date. The second group only had a deadline statement in the third and final email contact. The third group was informed of the deadline the second and third email contacts, and the fourth group was informed of the closing date in all three email contacts. All email contacts in Experiment II included a personal salutation, originated from a personal source email address, and were signed by an individual with the high authority title in the low authority office.

Results

Experiment I

The overall click to URL rate for Experiment I was 20.3% and the response rate for this experiment was 13.6%. Main effects for personalization of the email salutation, personalization of the email source address, authority of signatory, and authority of requesting office on click to URL and survey response rates were tested by comparing the two rates within each independent measure using a binominal z-test. As seen in Table 3, the click to URL and survey response rates did not differ significantly (at $p<.05$) across each independent measure. That is, URL click rates and survey response rates did not differ if the email salutation was personal or impersonal, if the email came from a personal or impersonal email address, if the authority of the signatory was high or low, or if the request came from an office high or low in authority.

Table 3. Experiment I: Main Effects of Personalization and Authority Conditions

Experimental group		Click to URL rates (%)	Response rates (%)
Salutation	Impersonal	19.7	12.9
	Personal	20.9	14.3
	p value	.16	.06
Source address	Impersonal	19.6	13.0
	Personal	21.1	14.2
	p value	.08	.10
Authority of signatory	Low	20.8	13.9
	High	19.9	13.3
	p value	.30	.42
Authority of requesting office	Low	20.1	13.5
	High	20.5	13.7
	p value	.64	.79
Total		20.3	13.6

Note: p values are from binomial z-tests.

Table 4 presents a more detailed breakdown of click to URL and response rates, and allows examination of the interactions between the four independent measures. To test for the cumulative effect of the four aspects of the email contact used in this experiment, the click to URL and response rates from the emails that included the most personalization and authority (personal salutation, personal source address, high title of signatory, and high familiarity with requesting) were compared to the control condition, emails with no personalization and the lowest level of authority (impersonal salutation, impersonal source address, low title of signatory, and low familiarity with requesting). Email contact including personalization originating from an office of high authority and individual of high authority did not significantly increase click to URL rates (20.7% vs. 22.3%, respectively) or survey response rates (13.1% vs. 15.5%, respectively) above the control. In addition, a logistic regression analysis testing for interaction effects did not indicate any statistically significant differences.

Table 4. Experiment I: Interactions of Personalization and Authority Conditions (%)

		Authority of requesting office: low		Authority of requesting office: high	
		Source email address: office	Source email address: person	Source email address: office	Source email address: person
Click to URL rates					
Title of signatory: low	Salutation: impersonal	20.7	20.2	20.0	21.5
	Salutation: personal	20.3	21.0	19.8	22.5
Title of signatory: high	Salutation: impersonal	19.4	19.7	15.6	20.8
	Salutation: personal	19.3	20.5	21.4	22.3
Response rates					
Title of signatory: low	Salutation: impersonal	13.1	12.3	13.4	15.5
	Salutation: personal	13.7	14.4	12.8	16.1
Title of signatory: high	Salutation: impersonal	13.8	13.1	9.6	12.7
	Salutation: personal	14.1	13.7	13.7	15.5

Note: Emails in Experiment I had no mention of selectivity or survey deadline.

These null findings further substantiate the results from the comparisons within the main effects: personalization of contact and authority of survey administrator did not affect response rates. That is, the techniques thought to increase response rates in traditional non-electronic surveys did not translate to our administration of a web-based survey using email contacts.

Experiment II

The overall click to URL rate for Experiment II was 23.7%, and the survey response rate was 17.4%. Click to URL and response rates are presented in Table 5. Since the click and response rates for the three levels of the deadline condition (deadline in all emails, deadline in the first and second emails, and deadline in the third email) were very similar within each selectivity condition, we combined the groups to form a general 'deadline' condition (i.e., anyone who received the statement about the survey closing date). The rates for the general deadline group are listed in the bottom of Table 5. This general deadline group was examined to test the overall effects of disclosing a survey deadline to potential respondents.

Main effects for statements of deadline and selectivity were tested by separately comparing the click to URL and response rates across the two levels of each independent measure. In addition, the effects of the deadline by selectivity interaction were tested by comparing rates when email contacts included both a deadline and a selectivity statement to the rates of those in the control condition (neither statement included in the three contact emails).

Main effects were found for both selectivity and deadline statements. Participants sent email messages that included a deadline were more likely to click on the URL ($z = 2.63, p<.01$) than participants who were not inform of a deadline. The benefit afforded by the inclusion of the deadline was a 3.8% increase in the rate of clicking to the URL (24.7% vs. 20.9%). When the email contact included a statement of selectivity, participants were more likely to click on the URL ($z = 4.70, p<.01$) and respond to the survey ($z = 5.12, p<.01$) than participants who were not informed that they were part of a select group. The selectivity statement increased the rate of clicking to the URL by 6.1% (26.8% vs. 20.7%), and increased survey response by 5.9% (20.3% vs. 14.4%).

As seen in Table 5, while the main effects for both deadline and selectivity increase click to URL and response rates, when we compare the table marginals the increase rates achieved by the inclusion of only one main effect statement at a time a different picture emerges. That is, inclusion of a deadline statement only (and not a selectivity statement) increases click to URL rates by only 0.2% above the control (20.7% vs. 20.5%, n.s.), and response rates by only 0.9% above the control (14.6% vs. 13.7%, n.s.). Similarly, inclusion of a selectivity statement only (and not a deadline statement) increases click to URL rates by only 0.8% above the control (21.3% vs. 20.5%, n.s). The inclusion of only a selectivity statement increased response rates by 3.8% above the control (17.5% vs. 13.7%). This difference, while moderate in magnitude, did not achieve statistical significance.

Table 5. Experiment II: Effect of Statements of Scarcity (%)

	Click to URL rates			Response rates		
	Selectivity statement:			Selectivity statement:		
Deadline statement:	No	Yes	Total	No	Yes	Total
No deadline	20.5	21.3	20.9	13.7	17.5	15.7
Deadline in email 3	21.2	29.9	25.4	15.7	21.4	18.5
Deadline in emails 2 & 3	19.7	29.7	24.7	12.5	20.8	16.6
Deadline in emails 1,2 & 3	21.3	26.7	24.0	15.6	21.9	18.7
Total	20.7	26.8		14.4	20.3	
No deadline	20.5	21.3	20.9	13.7	17.5	15.7
Deadline in at least one email	20.7	28.8	24.7	14.6	21.3	17.9
Total	20.7	26.8	*23.7*	14.4	20.3	*17.5*

Note: Emails in Experiment II had a personal source email address, used the participant's name in the salutation, were from the low authority office, and the title of signatory was high.

Thus, the main effects reported earlier appear to be driven primarily by the interaction between the deadline and selectivity statements. And indeed, the inclusion of, both, a deadline and a selectivity statement significantly increased click to the URL ($z = 3.98$, $p<.01$) and response rates ($z = 4.20$, $p<.01$) above the control, with increases of 8.3% (28.8% vs. 20.5%), and 7.6% in click and response rates (21.3% vs. 13.7%), respectively.

Deadline and selectivity statements both relay the scarcity of the opportunity for survey participation to potential respondents. That is, both statements tap the same construct, and it may be that, on their own, these statements do not relay sufficient information about the scarcity of participation. When paired together, however, they may provide enough information to surpass the critical threshold required for the potential respondent to realize the scarcity of their opportunity to participate, and thus, significantly increase click to URL and response rates.

Conclusion

Using a web survey sent to over 12,000 high school students, we investigated the impact of altering the email contact to understand the effects of personalization, authority, and scarcity on web survey response. Contrary to the literature on paper surveys, personalization of the email contact, whether through personalized greetings or a personal email address, appears to have little impact on response rates. This finding calls into question Dillman's (2000, p.368) recommendation concerning the use of personal emails versus emails coming from a listserv address. Authority of the survey sponsor also appears to have minimal impact, although in this case the null finding could be due to the fact that both survey sponsors were from a university. The low response rate to the survey is another possible cause for this result.

Statements of scarcity, however, did have a positive impact on response rates. Inclusion of a statement telling the respondent they had been selected as part of a small group to participate, and the inclusion of a deadline when the survey website would be shut down, both raised response rates almost eight percentage points. While not overwhelming, such an increase is quite respectable given that it was the result of only slight changes to the email message.

The most interesting question these results raise is, why does personalization of the contact appear to have no effect on the probability of survey response? The answer, we believe, lies in the nature of emails and today's Internet. Spam has become an ubiquitous feature of the web, and most people are used to receiving countless unwanted emails from marketers.

More importantly, alterations to emails are easily made and spammers take advantage of this. Spammers can easily alter the email address of origin, the signature of the sender, as well as use bulk email merging to personalize salutations. Recipients are quite used to receiving such spam emails and thus discount such alterations. Simply put, because these features are so easy to change, any such feature is no longer credible with email users. But it is still difficult to change the credibility of the message itself. Messages of scarcity are therefore more believable, especially when coming from an institution of higher education, and are more effective in raising response rates.

As spam continues to increases, annoyance with unsolicited emails will increase. Survey researchers using the Internet will continually have to refine their techniques in order to achieve a good response rate. It will become increasingly important for survey researchers to distinguish themselves from run of the mill spammers, and to do so in creative ways. For example, one experiment in a postal survey found that inclusion of a statement that reminders will be sent to nonrespondents increased the response rates by 10 percentage points (Green 1996). Such statements might prove even more effective in web surveys, as email users grow increasingly frustrated with unwanted emails. More research is needed into the effect of contact type on web survey response rates, and such research must be continually updated as use of the Internet changes.

References

Bosnjak, Michael, and Tracy L. Tuten. 2001. Classifying response behaviors in web-based surveys. *Journal of Computer Mediated Communication* 6 (3):1-14.

Bradley, Nigel. 1999. Sampling for Internet surveys: An examination of respondent selection for Internet research. *International Journal of Market Research* 41 (4):387-395.

Cho, Hyunyi, and Robert LaRose. 1999. Privacy issues in Internet surveys. *Social Science Computer Review* 17 (4):421-434.

Church, A.H. 1993. Estimating the effect of incentives on mail survey response rates: a meta-analysis. *Public Opinion Quarterly* 57:62-67.

Cobanoglu, Cihan, Bill Warde, and Patrick J. Moreo. 2001. A comparison of mail, fax and web-based survey methods. *International Journal of Market Research* 43 (4):441-452.

Cook, Colleen, Fred Heath, and Russel L. Thompson. 2000. A meta-analysis of response rates in web- or Internet-based surveys. *Educational and Psychological Measurement* 60 (6):821-836.

Couper, Mick P. 2000. Web surveys: A review of issues and approaches. *Public Opinion Quarterly* 64:464-494.

Couper, Mick P., Michael W. Traugott, and Mark J. Lamias. 2001. Web survey design and administration. *Public Opinion Quarterly* 65:250-253.

Crawford, Scott D., Mick P. Couper, and Mark J. Lamias. 2001. Web surveys: perceptions of burden. *Social Science Computer Review* 19 (2):146-162.

Dillman, Don A. 1991. The design and administration of mail surveys. *Annual Review of Sociology* 17:225-249.

Dillman, Don A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Dillman, Don A., and Dennis K. Bowker. 2001. The web questionnaire challenge to survey methodologists. In *Dimensions of Internet Science*, edited by U.-D. Reips and M. Bosnjak. Lengerich: Pabst Science Publishers.

Dillman, Don A Robert D. Tortora, Jon Conradt, and Dennis K. Bowker. 1998. Influence of plain vs. fancy design on response rates for web surveys. Paper read at Joint Statistical Meetings, at Dallas, TX.

Dommeyer, Curt J., and Eleanor Moriarty. 2000. Comparing two forms of an email survey: embedded vs. attached. *International Journal of Market Research* 42 (1):39-50.

Fox, Richard J., Melvin R. Crask, and Jonghoon Kim. 1988. Mail survey response rate: A meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly* 52 (4):467-491.

Goyder, John. 1987. *The Silent Minority: Nonrespondents in Sample Surveys*. Boulder: Westview Press.

Green, Josephine M. 1996. Warning that reminders will be sent increased response rate. *Quality and Quantity* 30 (4):449-450.

Groves, Robert M., Robert B. Cialdini, and Mick P. Couper. 1992. Understanding the decision to participate in a survey. *Public Opinion Quarterly* 56:475-495.

Heberlein, Thomas A., and Robert Baumgartner. 1978. Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review* 43 (4):447-462.

Heerwegh, Dirk, and Geert Loosveldt. 2002. Web surveys: The effect of controlling access using PIN numbers. *Social Science Computer Review* 20 (1):10-21.

Kiesler, Sara, and Lee S. Sproull. 1986. Response effects in an electronic survey. *Public Opinion Quarterly* 50 (3):402-413.

Kittleson, Mark J. 1997. Determining effective follow-up of e-mail surveys. *American Journal of Health Behavior* 21 (3):193-196.

Manfreda, Katja Lozar, Zenel Batagelj, and Vasja Vehovar. 2002. Design of web survey questionnaires: Three basic experiments. *Journal of Computer Mediated Communication* 7 (3):1-32.

Mehta, Raj, and Eugene Sivadas. 1995. Comparing response rates and response content in mail versus electronic mail surveys. *Journal of the Market Research Society* 37 (4):429-439.

Pealer, Lisa N., Robert M. Weiler, R. Morgan Pigg Jr., David Miller, and Steve M. Dorman. 2001. The feasibility of a web-based surveillance system to collect health risk data from college students. *Health Education & Behavior* 28 (5):547-559.

Presser, Stanley, Johnny Blair, and Timothy Triplett. 1992. Survey sponsorship, response rates, and response effects. *Social Science Quarterly* 73 (3):699-702.

Roberts, Robert E., Owen F. McCrory, and Ronald N. Forthofer. 1978. Further evidence on using a deadline to stimulate responses to a mail survey. *Public Opinion Quarterly* 42 (3):407-410.

Schaefer, David R., and Don A. Dillman. 1998. Development of a standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly* 62 (3):378-397.

Shannon, David M., and Carol C. Bradshaw. 2002. A comparison of response rate, response time, and costs of mail and electronic surveys. *Journal of Experimental Education* 70 (2):179-192.

Shannon, David M., Todd E. Johnson, Shelby Searcy, and Alan Lott. 2001. Using electronic surveys: Advice from survey professionals. *Practical Assessment, Research & Evaluation* 8 (1):1-14.

Sheehan, Kim. 2001. Email survey response rates: A review. *Journal of Computer Mediated Communication* 6 (2):1-20.

Sheehan, Kim Bartel, and Mariea Grubbs Hoy. 1999. Using e-mail to survey Internet users in the United States: Methodology and assessment. *Journal of Computer Mediated Communication* 4 (3):1-25.

Sheehan, Kim Bartel, and Sally J. McMillan. 1999. Response variation in email surveys: An exploration. *Journal of Advertising Research* 39 (4):45-54.

Sills, Stephen J., and Chunyan Song. 2002. Innovations in survey research: An application of web-based surveys. *Social Science Computer Review* 20 (1):22-20.

Smith, Tom W. 1995. Trends in nonresponse rates. *International Journal of Public Opinion Research* 7:157-171.

Steeh, Charlotte G. 1981. Trends in nonresponse rates, 1952-1979. *Public Opinion Quarterly* 59:66-77.

Tourangeau, Roger, Mick P. Couper, and Darby M. Steiger. 2001. Social presence in web surveys. Paper read at Federal Committee on Statistical Methodology Research Conference, at Arlington, VA.

Tynan, Daniel. 2002. Spam, Inc. *PC World* 20 (8): 1-5.

Vehovar, Vasja, Zenel Batagelj, Katja Lozar Manfreda, and Metka Zaletel. 2002. Nonresponse in web surveys. In *Survey Nonresponse*, edited by R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little. New Yrok: John Wiley & Sons.

Warriner, Keith, John Goyder, Heidi Gjertsen, Paula Hohner, and Kathleen McSpurren. 1996. Charities, no; lotteries, no; cash, yes. *Public Opinion Quarterly* 60:542-562.

Weible, Rick, and John Wallace. 1998. Cyber research: The impact of the Internet on data collection. *Marketing Research* 10 (3):19-23.

Willimack, Diane E., Howard Schuman, Beth-Ellen Pennell, and James M. Lepkowski. 1995. Effects of a prepaid nonmonetary incentive on response rates and response quality in a face-to-face survey. *Public Opinion Quarterly* 59:78-92.

Witte, James C., Lisa M. Amoroso, and Philip E. N. Howard. 2000. Method and representation in Internet-based survey tools: mobility, community and cultural identity in Survey2000. *Social Science Computer Review* 18 (2):179-195.

Yammarino, Francis J., Steven J. Skinner, and Terry L. Childers. 1991. Understanding mail survey response behavior: A meta-analysis. *Public Opinion Quarterly* 55 (4):613-639.

Yu, Julie, and Harris Cooper. 1983. A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research* 20:36-44.

Yun, Gi Woong, and Craig W. Trumbo. 2000. Comparative response to a survey executed by post, email and web form. *Journal of Computer Mediated Communication* 6 (1):1-25.

Zhang, Yin. 2000. Using the Internet for survey research: A case study. *Journal of the American Society for Information Science* 51 (1):57-68.

STATISTICAL METHODS FOR PREDICTING YIELD: A COMPARISON OF THE ACCURACY OF A LOGISTIC REGRESSION, DECISION TREE, NEURAL NETWORK AND BOOSTED LOGISTIC REGRESSION

Richard J. Reeves
Senior Research and Planning Associate
Office of Institutional Research and Planning, Cornell University

Martin T. Wells
Professor
Social Statistics Department, Cornell University

Introduction

Hossler (1984) defines enrollment management "as a process or an activity that influences the size, the shape, and the characteristics of a student body by directing institutional efforts in marketing, recruitment, and admissions, as well as pricing and financial aid." One of the primary functions of enrollment management for many institutions is controlling the size of the student body. A student body that is larger than anticipated can put excessive demand on university resources, such as housing, advising, dining, and class space, while a class smaller than anticipated can lead to fiscal shortfalls. For many institutions, incoming freshmen are a primary source of enrolled students. Setting a target for the number of students to admit to an institution requires that decision makers are able to accurately predict the proportion of accepted applicants that will enroll (yield). One of the ways institutions can limit its exposure to fluctuations in yield and better control how many incoming freshmen will attend in the fall is by taking a portion of its class through an Early Decision program that upon admission binds the student to attend that institution.

The increasing use of early decision in the admission process (Hooker-Haring, 1998) is indicative of a university's need to predict applicant behavior. When a student applies for early decision, that student agrees to attend that institution, if admitted. Typically, over 95% of the students admitted in early decision enroll as full-time students in the fall. The other 5% withdraw their commitment to the institution and often lose some or all of a deposit. Over the last 10 years, some highly selective universities increased the use of early decision from 10-20% of an entering class to 30-50% of an entering class (Hooker-Haring, 1998). Many institutions do not have the luxury of admitting a large portion of next year's class using early decision. Often the students who apply early decision are, on average, less qualified than are regular decision candidates. So institutions must balance high academic standards with the necessity to manage enrollment.

Bruggink and Gambhir (1996) created two models to assist with the admissions process. The first model predicted the likelihood that an applicant would be accepted and the second predicted the likelihood that an accept would enroll. While institutions attempt to admit the most qualified and diverse population with respect to institutional goals, students are focused on improving their qualifications to gain the admissions offer with the greatest perceived benefit. Thus the students who succeed in improving their qualifications are more likely to be admitted to the institutions where they apply. Meanwhile there is great overlap in institutional priorities within niches. The top admits at a given institution are least likely to enroll at any given institution. Consequently, it is the very process of selection that creates a population of admitted students that complicates predicting the proportion of students who will deposit given an accept decision (yield).

Bruggink and Gambhir (1996) reported a comprehensive list of independent variables that are appropriate to predict the probability of enrollment given an accept decision. The results of Bruggink and Gambhir (1996) highlighted academic factors (SAT scores, class rank, etc.), personal factors (major of interest, gender, ethnicity, etc.), extracurricular activities (leadership, community service, creativity, etc.), and special factors (child or relative of an alumnus, type of high school coming from, etc.) as components of both a selection model and a yield model. A logistic equation (regression) was used to determine the significance of the variables that went into the yield model. Given a list of acceptable predictors, a logical next step in the progression of research is to explore the accuracy of different statistical methods in predicting yield using a defined set of variables.

Two relatively new statistical methods that are appropriate to predict yield are decision trees (DT) and artificial neural networks (ANN). DT and ANN (sometimes referred to as feed forward neural networks) may be considered non-linear data mining statistics. Data mining is the process of searching for patterns in data without articulating a specific hypothesis. Data mining differs from traditional statistics in that there is not a significance test associated with an individual predictor (like a t-value or standardized Beta weight). Rather the model utilizes a predictor based on its ability to predict and does not necessarily assign a probability distribution or significance level to the predictor. Canned statistical packages like SAS Enterprise Miner (SAS EM) and SPSS Clementine allow a researcher to implement DT and ANN in a point and click environment. The mainstream availability of tools for DT and ANN permit an exploration of their utility in predicting yield.

The accuracy of a logistic regression, a decision tree, and a neural network are compared in predicting yield. All of these models are run from the SAS System Enterprise Miner module version 4.1. Enterprise Miner has a graphical user interface (GUI) that makes using different statistical methods convenient. This paper outlines the methods used to create the models and report the accuracy of modeling techniques in predicting the enrollment. A primary consideration in creating the models was the level of expertise required to create the models.

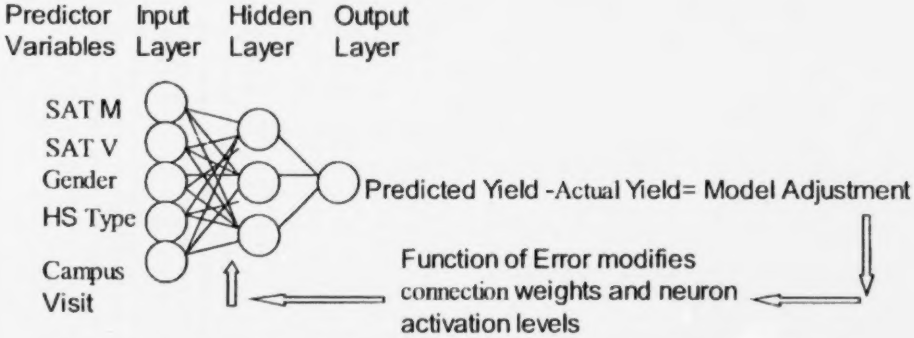
Statistical Methods

Logistic Regression is a special case of multiple regression and is designed to allow prediction of group membership when the dependent variable is dichotomous or categorical. The statistic is relatively free of restrictions and has the capacity to mix all types of predictors (continuous, discrete, and dichotomous) and interactions in the same model. (Tabachnick & Fidell, 1996). The history and use of logistic regression in higher education from 1988 to 2002 is well covered by Peng, So, Stage, & St. John (2002). Peng et al. report that there has been an increase in publications utilizing logistic regression and for this paper it can be considered the base (control) statistical method of choice for predicting yield as opposed to the newer statistical techniques like Artificial Neural Networks (ANN).

If we consider the decision to enroll as a consumer decision we may hypothesize from a statement made by West, Brockett, & Golden (1997) that ANN would be a superior prediction technique, "In the context of modeling consumer judgment and decision making, for example, Neural Network Models can offer significant improvement over traditional statistical methods because of their ability to capture nonlinear relationships associated with the use of noncompensatory decision rules." By design, ANN evaluates interaction effects and non-linear relationships without additional programming or modifying of variables. Gonzales and DesJardins (2001), explored the use of ANN to predict application behavior (likelihood to apply for admission) and found ANN to be an effective tool in predicting applicant behavior.

ANN are based on the known anatomy of mammalian brains. In general ANN comprise 3-4 layers. The first (left most) layer is the input layer and represents the predictor data that appear in the model. The final (right most) layer is the output layer and is where the final computed model prediction will appear. With any ANN, the model will measure the error of its predictions in the training dataset and adjust the model weights to become more accurate. The method of adjusting those weights is referred to as a learning algorithm and most packages will offer six or more of these learning algorithms. When training an ANN, the model will also track how accurately it is predicting a verification dataset. Since the model is not adjusting itself based on the verification dataset this accuracy level is a good indication of how well the model will predict on a novel dataset (the testing or scoring set).

Diagram 1: Representation of an ANN

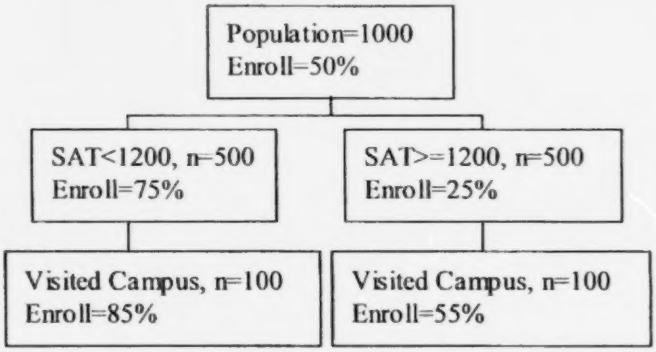

Predictor Variables
Input Layer
Hidden Layer
Output Layer
SAT M
SAT V
Gender
HS Type
Campus Visit
Predicted Yield -Actual Yield= Model Adjustment
Function of Error modifies connection weights and neuron activation levels

The number of neurons in the hidden layers (architecture of the ANN) can drastically affect the ANN's ability to generalize to novel data. If there are too many neurons in the hidden layers, the model may eventually 'memorize' the input dataset and predict outcomes with perfect accuracy or the model may become stuck in a local minimum and discontinue learning the data. If there are too few neurons, the model will not be able to accommodate the complex patterns in the data. While the architecture and learning algorithm are most often discussed when generating a model there are many other adjustments that can be made to the model that will affect its accuracy. A drawback to ANN is that they can be complex, difficult to interpret, and optimizing their accuracy is more intricate and involved than many other statistics. In contrast, decision trees provide a simpler nonlinear data mining technique.

Decision trees (DT) are a statistical technique that produces a model that can be represented with interpretable logic statements (SAS Enterprise Miner Help File, 2000). While DT are convenient classifiers of data, there is difficulty in the construction of a DT so that the hierarchy of the logic statements make sense from a policy or management standpoint (Sporgar et al., 2001).

An example of a hypothetical logic statement is: 1) Admitted students with an SAT total of 1200 or more are 25% likely to enroll; if they attended a visited campus as well, they are 55% likely to enroll. While the difference may be subtle, when discussing a topic like yield, decision makers will often want to discuss differences with respect to SAT score and not with respect to some nominal descriptor variable. Forcing a hierarchy of predictors with interval predictors like SAT coming first and other categorical predictors like control of high school coming later is not readily possible in SAS EM.

Diagram 2: Representation of a DT



SAS Enterprise miner also allows for a model to be boosted. Freund and Schapire (1999) describe Boosting as a general and probably effective method of producing a very accurate prediction rule by combining rough and moderately inaccurate rules of thumb to make a more accurate and sophisticated model. In practice, boosting re-weights each training observation based on the error of the prediction made. Misclassified cases are re-weighted to have a greater weight in the model, while correctly classified weights are re-weighted to have a lesser presence in subsequent iterations of the model (SAS, 2000). In theory, after boosting a model many times, the combined effects of those (posterior distributions) models create a more accurate final model than was originally calculated on the first iteration. The logistic regression will be boosted and its accuracy measured.

Methodology

Three years of admission data were collected for regular decision (non-early decision) fall term freshmen from a selective research I institution in the north east. A data partition node was used to separate the three sequential years of applicant data into: a training dataset (current year), a validation dataset (previous year), and testing dataset (future year). The training dataset will be used by each statistical method to generate the predictive model. The Decision Tree and the Neural Networks will use the validation dataset during model creation to the modeling process to evaluate the accuracy of the model on a novel dataset. The completed models will be scored against the (novel) testing dataset and measured for accuracy in predicting yield. This method of partitioning the data allows for a model to be created using the most recent year and verified using the previous year and then tested using a novel dataset as would happen in practice.

The variables selected to predict yield are shown in Table 1. The variables selected for this process were evaluated and selected based on availability, accuracy, and relevance to the college choice model outlined in Bruggink and Gambhir (1996). No financial aid information was used in this model, as those data were not available for all three years. State was collapsed into the top 13 states with the most accepted students and a 14th value for admits from all other states. CONTOTAL represents the total number of contacts the institution and student had prior to March 1 of the academic year of application. This allows for some prospect information to be included in the model, while giving a point in time when the yield model can be fully implemented. Similarly, CONSTUD represents the total number of student-initiated contacts with the university by the same date and CP_VIS indicates if a student has visited the campus at least 1 time prior to March 1.

Table 1: Variables Used in Models

Name	Role	Description
CITIZEN	Input	Student is a US Citizen (Y/N)
CONSTUD	Input	Number of contacts initiated by the student
CONTOTAL	Input	Total number of contacts with student
CP_VIS	Input	Student Visited Campus (Y/N)
ETHNO	Input	Student identified Ethnicity
FSC	Input	Parent works at Institution
HSATM	Input	Highest submitted SAT Math Score
HSATV	Input	Highest submitted SAT Verbal Score
LEGACY	Input	Student has a parent or grand parent who attended the institution
PROGRAM	Input	Academic program student is applying for
SEX	Input	Gender
STATE	Input	Home State
ENROLL	Target	Student Decided to Enroll (Y/N)

The dependent measure (enroll) is based on the admitted student's intent to enroll at the end of the admissions cycle rather than actual fall enrollment. The reason for this is that there are a number of admitted students who deposit in the spring and never enroll in the fall. The number of these no-show depositors varies from year to year (1%-3% of the incoming class of the subject institution) but does not appear to be related to the variables included in the model. This is an important distinction, because it means that when interpreting the accuracy of the models, the reader should understand that the accuracy is for a model set to a date in May and not in September.

For the logistic regression, the variables listed in Table 1 will be entered into the model, then the variable with the highest p-value that is also greater than .2 is removed from the model, this process will be repeated until all variables included in the model had a probability of less than .2. Then all possible interactions for the remaining variables were entered into the model. The interaction with the highest p-value greater than .2 will be removed and the model re-run. The set of variables and interactions included in the final model are shown in Table 2 with their respective Wald statistic. Table 3 shows that the entire model is significant.

Table 2: Type III Analysis of Effects for Logistic Regression

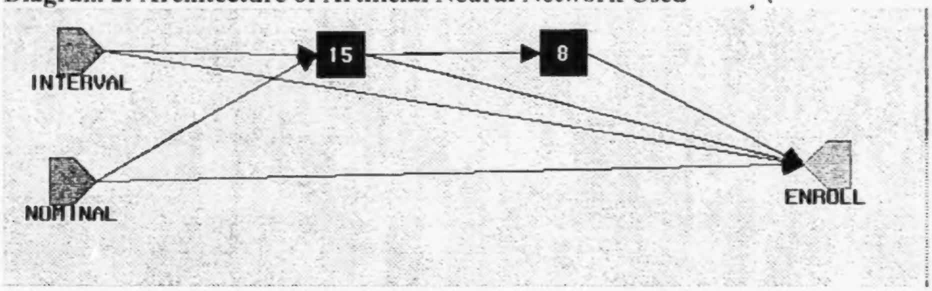
Variable	DF	Wald Chi-Square	Pr >Chi Square
CONSTUD	1	1.8102	0.1785
CONSTUD*LEGACY	1	6.9122	0.0086
CONSTUD*SEX	1	6.471	0.011
CONTOTAL	1	1.9526	0.1623
CONTOTAL*ETHNO	6	15.4289	0.0172
CONTOTAL*HI_SATM	1	2.7206	0.0991
CONTOTAL*LEGACY	1	4.4148	0.0356
CP_VIS	1	7.0593	0.0079
ETHNO	6	17.0408	0.0091
ETHNO*SEX	6	16.6189	0.0108
HI_SATM	1	2.144	0.1431
HI_SATM*HI_SATV	1	29.0636	<.0001
HI_SATV	1	0.5211	0.4704
HI_SATV*ETHNO	6	12.417	0.0533
LEGACY	1	1.5535	0.2126
PROGRAM	6	65.545	<.0001
PROGRAM*CP_VIS	6	16.4679	0.0115
PROGRAM*SEX	6	9.5539	0.1447
SEX	1	8.306	0.004
STATE	13	57.8018	<.0001
STATE*SEX	13	30.1383	0.0045

Table 3: Testing the Global Null Hypothesis that Beta=0.

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
-2 LOG L	5924	5208	7.15.5 with 80 DF (p<.0001)

The neural network node in Enterprise Miner has an advanced level of control that allows for design of the neural network, activation functions of neurons, learning algorithm for training and preliminary weights, and many other important factors to the design of an ANN. To establish beginning model weights 100 preliminary runs were made using the Quasi-Newton training technique to establish a starting point with the ANN. The training technique used in model creation is the RPROP technique. The RPROP technique was chosen because it is the most stable of all the 'prop' techniques available and rarely requires any tuning (SAS, 2000). For this model the first hidden layer had 15 neurons and the second had 8. The input layer and each hidden layer were directly attached to the output layer, as well as in feed forward order (input - hidden 1- hidden 2) as shown in diagram 1.

Diagram 2: Architecture of Artificial Neural Network Used



The DT was set to imitate the performance of a C5.0 DT as is covered in the SAS manual (SAS, 2000). Minimum counts in terminal nodes were set to 15 and to 40 for minimum number of observations for a search. The splitting criteria employed were set to entropy reduction. The final models for the logistic regression and the DT were boosted. The boosting sequence was run for 100 iterations and the models were combined.

Results

The accuracy of the four different statistical methods is shown in Table 4. When estimating accuracy, the sum of the predicted probabilities was used. This is an alternative to the concordance tables which force the predicted probabilities into either a predicted enroll or not enroll. The DT was the most accurate in predicting yield. The boosted logistic regression was the least accurate. The logistic regression under predicted enrollment by 26 and artificial neural network over predicted enrollment by 36.

Table 4: Accuracy of Statistical Models on Test Year Data

Statistical Method	Predicted Yield	Actual Yield	Predicted-Actual	Admits in Test Year	Predicted Enrollment Error
Logistic Regression	37.01%	37.56%	-0.55%	4,840	-26
Artificial Neural Network	38.30%	37.56%	0.74%	4,840	36
Decision Tree	38.05%	37.56%	0.49%	4,840	24
Boosted Logistic Regression	48.09%	37.56%	10.53%	4,840	510

Discussion

The DT provided the most accurate prediction in this experiment (over-predicted enrollment by 24 students or just over 1 percent of the enrolling class). The logistic regression was a close second by under-predicting the class by 26 students. The ANN network came in third with an over-prediction of enrollment by 36 students. We can see from these findings that all three of the statistical methods employed are viable options for predicting yield. However, given the complexity of ANN, it appears that the most pragmatic choices are the logistic regression and the DT.

The increase in error that occurred with the boosted logistic regression is surprising. The authors are unaware of any published papers citing an increase in error rate after using a boosting algorithm. This prompted a phone call to the SAS institute where a statistical support representative explained that this could happen with the algorithm employed by Enterprise Miner (a proprietary algorithm) when the dataset has a high level of noise or the model is accounting for a low level of variance. The data being used to predict yield does not account for a lot of variance. Essentially, the logistic regression model kept boosting noise and consequently increasing the error rate. For these reasons it seems reasonable to state that the boosting algorithm in SAS may not be appropriate for yield prediction. Retesting this finding is necessary particularly as the practice of boosting is relatively new and other specialized boosting algorithms are added to canned statistical programs.

It is also apparent from reviewing the training and testing data more closely that those two data sets are more similar with regard to yield prediction than the validation data set. This odd validation dataset could have misguided the development of the DT and ANN. The similarity between years is an important item to consider when using data mining techniques to create models.

For future research, it is recommended that a researcher determine the list of variables that will be used on the yield models and then amass 10 or more years of data. Calculate the yield model for each training year, validate on the previous year, and test on the year after and repeat the model creation for all but the first and last year of data available (because there would be no validation or testing dataset respectively). This empirical method will give an error amount for each year allowing for the calculation of mean error.

Evidenced by the low error rate, it can be concluded that Bruggink and Gambhir (1996) provided a sufficient list of variables that predict yield that were available for this study. Others repeating this method for their own use would be advised to include financial aid data in the model as well as any high school information that is stored in the data systems.

By far, the simplest model to create was the DT. The options and their effects are well documented. The ANN is a very complicated statistical procedure to maximize. Even for an experienced researcher the ANN is time consuming and processor intensive.

There are many combinations of architecture, learning algorithms, starting and stopping criteria, and activation functions that should be considered in creating a model.

The necessity to enroll a class on target will most likely be a priority for institutions for some time to come. With an ever growing set of statistical methods for researchers to employ, future research in predicting yield should include an evaluation of relevant statistical methods as well as continuing to consider appropriate variables for inclusion in the yield model.

Bibliography

Byers Gonzalez, J. M., DesJardins, S. L. (2001) Artificial Neural Networks: A New Approach for Predicting Application Behavior. Association for Institutional Research (AIR) 41st Annual Forum. Long Beach, California.

Bruggink, T.H., and Gambhir, V. (1996). Statistical models for college admission and enrollment: A case study for a selective liberal arts college. Research in Higher Education, Vol. 37 n2, pp221-240.

Enterprise Miner Reference Manual [SAS System Version 8.2; Enterprise Miner Version 4.1] (2000). Cary, NC: SAS Institute Inc.

Freund, Y. and Schapire, R. E. (1999). A Short Introduction to Boosting. Journal of Japanese Society for Artificial Intelligence, 14(5): 771-780, September, 1999.

Hooker-Haring, C. (1998). Playing the college admissions game. High-School-Magazine, v5, p32-37.

Hossler, D. R. (1984). Enrollment Management: An integrated approach. New York: The College Board, 1984.

Peng, C., So, T., Stage, F. K., St. John, E. P. (2002). The Use and Interpretation of Logistic Regression in Higher Education Journals: 1988-1999. Research in Higher Education, v43 (3): 259-293.

Sprogar M., Kokol P., Zorman M., Podgorelec V., Lhotska L., Klema J. (2001): Evolving Groups of Basic Decision Trees. Proceedings of the 14th IEEE Symposium on Computer-based Medical Systems CBMS (pp. 17-22). Bethesda, ML, USA.

Tabachnick, B. G. & Fidell, L. S. (1996). Using Multivariate Statistics, (3rd ed.). USA: HarperCollins College Publishers.

West, P. M., Brockett, P. L., and Golden, L. L. (1997). A comparative analysis of neural networks and statistical methods for predicting consumer choice. Marketing Science, 16 n 4, pp. 370-391.

PAPER VS. WEB: THE DIFFERENTIAL IMPACT ON RESPONSES OF MEN AND WOMEN

Heather S. Roscoe
Senior Research Analyst

Dawn Geronimo Terkla
Executive Director
Office of Institutional Research
Tufts University

Introduction

The Internet is quickly becoming an integral part of everyday life. More and more survey research (including that done in higher education and Institutional Research) is being conducted on the web, since web-based surveys seem to be beneficial in terms of cost, time, and accuracy (Underwood, Kim, & Matier, 2000). To date, there is little research on the validity and reliability of data obtained from web surveys as compared to paper and pencil surveys. Moreover, few have examined whether men and women react differently to varied survey formats, and if their response rates are affected by different formats. The minimal research that has been conducted has not yielded consistent results. Therefore, it is important to determine whether data obtained from web surveys are as valid and as reliable as those of paper surveys, if response rates for men and women differ between paper and web surveys, and if men and women are differentially affected by a change in survey format.

Only a handful of studies have investigated how respondents react to web surveys, particularly the issues of bias in responses compared to paper and pencil surveys, and differences in overall response rates. With respect to response rates, Underwood, Kim, and Matier (2000) and Cross, Thomson, and Daly (2001) both found that a mail (paper) survey had a notably higher response rate than an identical web survey. In addition, some research has indicated that females tend to respond in a greater proportion than males regardless of the method of survey administration (Underwood, Kim, & Matier, 2000), while other results suggest that web surveys yield a greater proportion of male than female respondents (Tomsic, Hendel, & Matross, 2000).

Other researchers have explored the issue of response bias or differences in responses based on method of survey administration. Bliven, Kaufman, and Spertus (2001) reported that data collected via the web were highly correlated with data obtained from the paper-and-pencil version of the survey. Epstein, Klinkenberg, Wiley, and McKinley (2001) randomly assigned undergraduate students to either a paper format group and a web format group. They found that the two versions of the survey elicited very similar results overall, but when responses were analyzed holding gender constant, the paper version of the survey received higher ratings than the web version. Baron and Siepmann (2000) found that responses to web surveys were very similar to the responses to paper surveys, but only if the formatting was exactly the same. Cross, Thomson, and Daly (2001) also found that web surveys evoked similar responses to paper surveys, with the greatest differences being in response to questions about automobile expenses, computer ownership, and marital status. Further, some studies have found that web surveys tended to receive significantly more favorable responses than paper surveys (Carini, Hayek, Ouimet, & Kuh, 2001; Tomsic, Hendel, & Matross, 2000). Research conducted by Underwood, Kim, and Matier (2000) suggested that web surveys elicited less favorable responses than did paper surveys. Some findings have indicated that participants tend to respond in a more honest fashion on the web (Turner & Rogers, 1998), while other findings suggested that respondents give more guarded/socially desirable responses on the web (Antons, Dilla, & Fultz, 1997).

Some research suggests that even in this day and age when many people (especially young adults) tend to be fairly familiar with current technology, female college students feel less confident, less positive, more anxious about using computer and Internet technologies, and tend to use the technologies less frequently than men (Jackson, Ervin, Gardner, & Schmitt, 2001; Mitra, Lenzmeier, Steffensmeier, Avon, Qu, & Hazen, 2000; Schumacher, & Morahan-Martin, 2001). In addition, women tend to use email and the Internet for communicating with others or educational assistance, while men use email and the Internet in order to gain information or for entertainment/leisure (Jackson et al., 2001; Weiser, 2000). After converting an annual residential life survey from paper-and-pencil to the web, Roscoe, Terkla and Dyer (2002) found that the change in survey administration method affected males and females differently (i.e. there were some statistically significant interactions between gender and survey format). According to the above findings, the experience of using computers and the Internet may be so different for men and women that their responses to web surveys may be biased by the experience. This might lead to a change in historical trends for each gender at the point of changeover from paper-and-pencil surveys to web surveys. This possibility warrants further investigation on the topic.

The primary objective of this paper is two-fold: 1) to investigate if there are differences in the responses of men and women that are dependent on the format of survey administered (web vs. traditional paper and pencil) at a Carnegie classified doctoral extensive university, and 2) to investigate if the overall response rates of men and women change as a result of different survey administration formats. In order to explore this area, two years of data from an Accepted Applicant web-based survey were compared to two years of data from the paper administrations of the same survey.

Methodology

This paper used a longitudinal research design, where the trend of a given general population was analyzed at different data-collection points. At this university, a number of surveys are administered annually. Using data from a survey that has been administered for over fifteen years, this paper explored 1) whether response rates were impacted by the delivery mechanism, and 2) whether the delivery mechanism (paper vs.

web) had a differential effect on male and female accepted applicants. The analyses were conducted using Univariate ANOVAs.

Applicants who are offered admission to the university were invited to respond to one of two Accepted Applicant surveys. One survey was sent to enrolling students (new student survey), and one to students who elected not to enroll (non-enrolling student survey). The majority of the questions on both surveys were identical. The questions asked, the ordering of questions, and format of questions within the survey instruments have remained fairly constant from one year to the next. For the web administrations, both survey instruments were shortened to facilitate online completion. The questions that remained were the same as those from previous years. For purposes of this analysis, data from the 1999 and 2000 paper surveys and the 2001 and 2002 web surveys were used. Only items that had been included for all four years were included in the analyses. For the past two years, the surveys were developed using software available from a private web surveying company and were published on their server. There have never been any incentives offered for completing the survey.

Results

The response rates to the new student web surveys were slightly lower than response rates of previous paper administrations, while curiously the non-enrolling student web surveys elicited a slightly higher response rate than that of previous paper surveys. See Table 1.

Table 1

Number of Respondents, Number of Potential Respondents, and Corresponding Response Rate

	Number of Respondents	Number of Potential Respondents	Response Rate
		New Student	
Paper	1,698	2,530	67.1%
Web	1,409	2,587	54.5%
		Non-Enrolling Student	
Paper	1,821	5,407	33.7%
Web	1,601	4,508	35.5%

The gender distribution of the respondents was fairly similar across survey formats. This was true for both the new student and non-enrolling surveys. See Table 2.

Table 2

Number of Respondents by Gender

	Male		Female		Total	
	n	%	n	%	n	%
			New Student			
Paper	748	44.7%	925	55.3%	1,673	100.0%
Web	497	40.3%	735	59.7%	1,232	100.0%
			Non-Enrolling Student			
Paper	725	40.3%	1,076	59.7%	1,801	100.0%
Web	557	40.4%	823	59.6%	1,380	100.0%

The distribution of respondents' ethnicities was similar between the two survey formats, and this was true for both the new student and non-enrolling student surveys. See Table 3. Some of the changes in the distribution over the years may be accounted for by the fact that, in general, the proportion of most of the minority groups at this university has been increasing.

Table 3

Number and Percent of Respondents by Ethnic Identity

	African American		Asian American		Hispanic American		Native American		White American		Total	
	n	%	n	%	n	%	n	%	n	%	n	%
							New Student					
Paper	112	7.9%	239	16.9%	121	8.6%	0	0.0%	939	66.5%	1,411	100.0%
Web	79	7.3%	168	15.6%	125	11.6%	8	0.7%	696	64.7%	1,076	100.0%
							Non-Enrolling Student					
Paper	110	7.4%	318	21.4%	135	9.1%	5	0.3%	917	61.8%	1,485	100.0%
Web	113	9.4%	252	21.1%	144	12.0%	5	0.4%	682	57.0%	1,196	100.0%

In addition, respondents to web surveys did not significantly differ from the respondents to paper surveys on a variety of demographic characteristics such as parental marital status, the type of area in which they live (rural, urban, suburban, etc.), whether and how long they have lived outside of the United States, and whether English was their native language. Therefore, since the respondents to the paper surveys seem to resemble the respondents of the web surveys so closely in terms of many demographic characteristics, it is unlikely that any statistically significant interactions between gender and survey format can be attributed to differences between the samples.

The overall patterns of responses were similar between the two methods of survey administration of both the new student and non-enrolling versions of the survey. However, when the results were analyzed using a series of Univariate ANOVAs, some main effects of survey format, some significant main effects of gender, and some significant interactions between the two emerged.

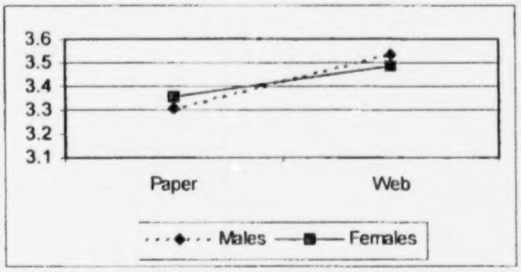
In general, females tended to give more positive ratings than males to various aspects of the university such as the diversity of the student body, campus appearance, various facilities, housing, student services, and internship and research opportunities. Moreover, females were more likely than males to report that they felt they would succeed at the university such as obtaining an internship, graduating with multiple majors and minors, that they would study abroad, participate in student protests or demonstrations, and become active in a cultural center. Females were also more likely to report that they would probably change majors, meet with their advisor often, and would need to obtain a job while attending college. Males were more likely than females to indicate that they were likely to graduate in the top half of their class, would participate in varsity and/or intramural athletics, and also expressed more positive feelings about the presence of fraternities and sororities on campus.

Overall, the transition from paper to web tended to elicit more positive ratings for characteristics of and opportunities provided by the university, and web respondents were also more likely than paper respondents to report that they were likely to succeed in various ways such as graduating with honors and they were more likely to partake of the opportunities on campus such as study abroad and becoming active at a cultural center. It is quite possible that some of these significant differences could very well be indicative of changes that have been made at the university in recent years.

As shown above, there were some consistent effects of gender and some consistent effects of survey format. However, the purpose of this paper was to explore whether there were any differential effects between males and females following a change in survey format from paper to web. Among the 108 survey items (69 from the new student survey and 39 from the non-enrolling student survey) that were investigated, only five variables were found to have significant interactions between survey format and gender, meaning that the transition from paper to web format affected the responses of men and women differently. Those variables in the new student survey with significant interactions between gender and survey format included the ratings of the quality of campus social life, the number of course offerings, and campus safety. See Charts 2 through 4. The variables in the non-enrolling student survey that were found to have a significant interaction between gender and survey format were the ratings of the quality of campus housing and the university's guidebook rankings. See Charts 5 and 6. For the most part, these variables deal with quality of life issues, which are areas that may be of more concern to females. The general pattern was for females to give more positive ratings than males during paper administrations, and for the ratings of both males and females to grow more positive from paper to web. However, for males the increase was such that it surpassed those for the females and thus the male respondents' ratings of these particular items were more positive.

Chart 2

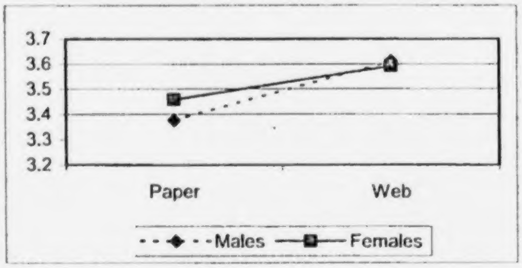
New Student: Mean Ratings of the Quality of Campus Social Life by Gender and Survey Format



Scale: 1 = Poor, 2 = Fair, 3 = Good, 4 = Excellent

Chart 3

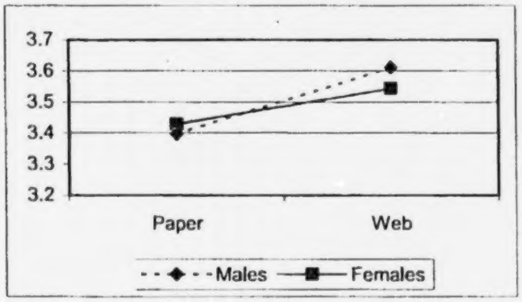
New Student: Mean Ratings of the Number of Course Offerings by Gender and Survey Format



Scale: 1 = Poor, 2 = Fair, 3 = Good, 4 = Excellent

Chart 4

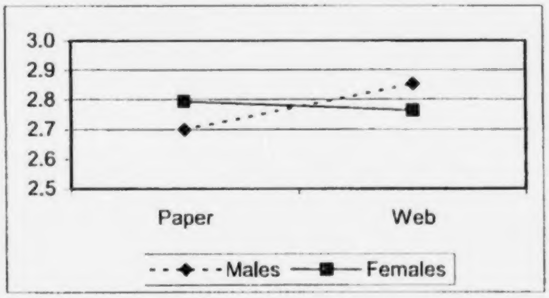
New Student: Mean Ratings of Campus Safety by Gender and Survey Format



Scale: 1 = Poor, 2 = Fair, 3 = Good, 4 = Excellent

Chart 5

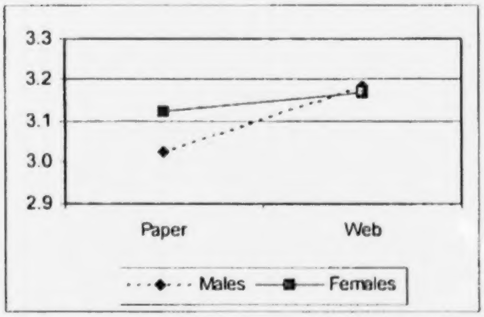
Non-Enrolling Student: Mean Ratings of Quality of Campus Housing by Gender and Survey Format



Scale: 1 = Poor, 2 = Fair, 3 = Good, 4 = Excellent

Chart 6

Non-Enrolling Student: Mean Ratings of Guidebook Rankings by Gender and Survey Format



Scale: 1 = Poor, 2 = Fair, 3 = Good, 4 = Excellent

Less than 5% of the items investigated revealed significant interactions between gender and survey format. In addition, those items that did experience a significant interaction were not largely affected in that the changes in mean ratings were relatively small and therefore do not seem to suggest a huge effect on either males or females by a switch in the format of the survey. A certain amount of fluctuation over time should be expected. The fact that the overall pattern of responses did not change drastically from the switch from paper to web surveys may indicate that the switch did not dramatically affect the responses of either males or females, that the status quo was not overly disrupted by the conversion.

Despite the fact that the change in survey format did not seem to differentially affect the results of males and females, curiosity led to a limited analysis of the use of computers and technology between the genders. The only questions that were asked all four years related to why the respondents did not apply to this university using an electronic application. Males and females tended to respond to these questions in a similar fashion, but with females tending to have very slightly less computer skills, less access to computers, less adequate computers, and less confidence about how the application would appear when it arrived at the admissions office than their male counterparts. However, females were more aware than males that there was an electronic option, and they were more likely than males to feel an electronic application was more impersonal. Perhaps women are slightly less confident than males in their ability to use computers and technology as some of the previous research in this area suggests, but this lack of confidence did not seem to translate into a decreased response rate of females after the conversion nor did it lead to an unexpected shift in pattern of responses to the questions on this particular survey at this university.

Discussion

Our analyses of responses to this survey over time did not reveal any drastic changes in response rates from both males and females, nor did they reveal a great number of items that males and females responded to differentially depending on survey format. Therefore, converting the survey format from paper to web has not seemed to greatly affect the responses to this survey at this university. However, this institution may not be representative of other institutions. In addition, the literature review highlights that not enough research has been done to reach a consensus about the effects of making a switch in survey format. Therefore, institutional researchers who are considering making a transition from paper to web surveys may want to consider their own population, and whether the change might affect their respondents differently than what has been experienced at this institution. Moreover, after converting from paper to web surveys, it might be useful to carefully monitor the response such as if new gender differences emerge, or if there are drastic changes in responses the year of the conversion. Another suggestion would be that, prior to embarking on web-based administration of surveys, it would be useful to obtain information regarding the current populations usage and related anxiety while using computers, the Internet, and technology. These findings might help inform the decision of whether it is "the right time" to switch from paper to web-based administrations.

References

Antons, C. M., Dilla, B. L., & Fultz, M. L. (1997). Assessing student attitudes: Computer versus pencil. Paper presented at the Annual Forum of the Association of Institutional Research (37th, Orlando, FL, May 18-21, 1997).

Baron, J., & Siepmann, M. (2000). Techniques for creating and using web questionnaires in research and teaching. In M. H. Birnbaum (Ed.), Psychological experiments on the internet (pp. 235-265). New York, NY: Academic Press.

Bliven, B. D., Kaufman, S. E., & Spertus, J. A. (2001). Electronic collection of health-related quality of life data: Validity, time, benefits, and patient preference. Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 10(1), 15-21.

Carini, R., Hayek, J. C., Ouimet, J. A., & Kuh, G. D. (2001). Student responses on web vs. paper surveys: A test of mode effects. Paper presented at the Annual Forum of the Association of Institutional Research (41st, Long Beach, CA, June 3–6, 2001).

Cross, J., Thomson, G., & Daly, R. (2001). Web vs. paper surveys: Lessons learned from a direct large-scale comparison. Paper presented at the Annual Forum of the Association of Institutional Research (41st, Long Beach, CA, June 3–6, 2001).

Epstein, J., Klinkenberg, W. D., Wiley, D., & McKinley, L. (2001). Insuring sample equivalence across internet and paper-and-pencil assessments. Computers in Human Behavior, 17(3), 339-346.

Jackson, L. A., Ervin, K. S., Gardner, P. D., & Schmitt, N. (2001). Gender and the internet: Women communicating and men searching. Sex Roles, 44(5-6), 363-379.

Mitra, A., Lenzmeier, S., Steffensmeier, T., Avon, R., Qu, N., & Hazen, M. (2000). Gender and computer use in an academic institution: Report from a longitudinal study. Journal of Educational Computing Research, 23(1), 67-84.

Roscoe, H. S., Terkla, D. G., & Dyer, J. A. (2002). Administering surveys on the web: Methodological issues II. Paper presented at the Annual Forum of the Association of Institutional Research (42nd, Toronto, ON, June 2-5, 2002).

Schumacher, P., & Morahan-Martin, J. (2001). Gender, internet and computer attitudes and experiences. Computers in Human Behavior, 17(1), 95-110.

Tomsic, M. L., Hendel, D. D., & Matross, R. P. (2000). A world wide web response to student satisfaction surveys: Comparisons using paper and internet formats. Paper presented at the Annual Forum of the Association of Institutional Research (40th, Cincinnati, OH, May 21-24, 2000).

Turner, C. F., Rogers, S. M., Lindberg, L. D., Pleck, J. H., Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. Science, 280, 867-873.

Underwood, D., Kim, H., & Matier, M. (2000). To mail or to web: Comparisons of survey response rates and respondent characteristics. Paper presented at the Annual Forum of the Association of Institutional Research (40th, Cincinnati, OH, May 21-24, 2000).

Weiser, E. B. (2000). Gender differences in internet use patterns and internet application preferences: A two-sample comparison. CyberPsychology & Behavior, 3(2), 167-177.

THE FREQUENCIES OF STUDENT ONLINE ACTIVITY AS PREDICTORS OF COURSE GRADE

Charles Secolsky
Director of Institutional Research & Planning

Cliff L. Wood
Vice President of Academic Affairs
County College of Morris

Introduction

Technology has changed the practice of distance learning in a very short time (Turoff, 1997). There are now a number of web-based testing and grading systems that keep track of student online activity. Wang and Newlin (2000) state that the monitoring of student activity in web-based instruction is a key component of the student-teacher relationship, because the instructor has no other way to gauge or judge student motivation or effort. "The usual set of cues such as facial expressions or even fidgeting that might be indicative of student confusion or disinterest is absent." (Wang and Newlin, 2000: p.142).

One instructional system that has gained in popularity is WebCT, which not only offers the instructor a means for teaching and grading online, but also allows the instructor to collect and report, for each student, three frequencies, in addition to the dates of the student's first and last access. The online system, WebCT, used at County College of Morris, collects and reports for each student: (1) the number of "hits" or number of times the student logged into the course, (2) items "read" or the number of content pages students read in HTML, and (3) items "posted" or the number of items students posted to the bulletin board. While these three frequencies are not evidence of learning as would be supplied by a test, they do indicate, especially in the case of items "posted," work evaluated by the instructor. This depends, of course, on the type of course in which the online instruction is used. In three web-based sections of psychological statistics, Wang and Newlin (2000), among other things, found that total online activity was predictive of final grades in the course. Kendall (2001) produced a matrix of grade level by percentages of items "read" and items "posted." She concluded that there was a positive correlation between achieving higher grades and frequency of accessing course notes and read bulletin board messages. Also Finnegan and Finnegan's findings (2002) suggest that regular use of WebCT seems to be related to student persistence in the course.

Following this tradition, the present study also looked at the relationships between the frequencies of the indicators of online performance and final grade in the course for five class sections. Relationships were examined for the five sections combined and for each individual section. Oncethe relationships were obtained, the question was asked as to why such a relationship should exist and how it would or would not manifest itself in actual practice. The same basic result was found for the frequencies of online activity as predictors, although, differences were found between different types of courses and different instructional uses of online activity.

Wang and Newlin (2000) reported a correlation of "hit" rate with class grade of .38. The present study found a correlation of .39, which seems like a fairly decent replication. But, if one were to look at the distribution more closely, one would notice that students with an extremely high number of hits would not necessarily get the highest grades. Why not? Perhaps, students who access the course too frequently may do so with the intention of only checking continually whether their postings received a response. Or, it may be that these students are new to the technology behind online instruction.

This complexity may explain why the findings are not exactly the same as those of Wang and Newlin (2000). To try to uncover the nature of this and other anomalies, case interviews were used after the statistical analysis to flesh out the meaningfulness of the statistics. This approach is the reverse order of what is typically done with case studies and statistical analysis. The case interviews are meant to reflect how online activity is used by instructors and query why it might be the case that the frequencies are correlated with final grade.

Methodology

Five sections of online courses were analyzed. The sections were New Media Technology (n=10 students), American Literature-Colonial to the Civil War (n=13 students), American Literature-Civil War to the 20th Century (n=15 students), Creative Writing (n=14 students), and English Composition (n=12 students). These 64 students received final grades. The frequencies of online activity (i.e., number of "hits," number of items "read," and number of items "posted") were analyzed in relation to final grade in the course for all 64 students taken together and separately by section. Pearson product-moment correlations were computed and multiple regressions were applied. It was conjectured that the more frequently students had these types of activities, the greater would be their participation in the course, the greater the effort, the higher the final grade. No attempt, however, was made to control for pre-existing student ability.

The regression analysis was performed on the entire group of 64 students and separately for each of the students in each of the sections studied. Two of the three instructors taught two courses each and a third instructor taught one course. Students with W grades were excluded as were students whose last access date was not near the end of the semester. Thus, students with grades of "F" due to lack of attendance were eliminated.

Prior to running the regressions and correlations, univariate statistics were computed to check for unreasonable values of skewness and kurtosis for any of the three frequency variables. With a skewness above 2.0 for items "posted," data were transformed using the square root function. In a reverse of standard methodological procedures, interviews were conducted with online instructors to enable an understanding of the statistical results. The interviews were intended to both corroborate the statistical evidence and explain difficult to explain conclusions from the analysis of the data alone.

Results

The means, standard deviations, minimums and maximums are presented below for each of the frequency variables.

Means, Standard Deviations and Minimums and Maximums
for Grade and Each of the Frequency Variables for the 64 Students*

Variable	Hits	Reads	Posts	Grade
Mean	499.1	282.4	26.5	2.5
St. Dev.	335.1	248.0	27.0	1.1
Minimum	49.0	1.00	1.00	0
Maximum	1513	896.0	123.0	4

* Grades are on a scale of 0 (F) to 4 (A).

All three frequency variables taken together as predictors of course grade yielded a significant regression model (p=.0042). Adjusted R Square = .156, however, the regression coefficients were not significant. The same was true for all pairwise entries of variables into the equation. However, when each of the variables was entered independently, the coefficients in all three cases were significant. The highest proportion of variation explained was attributable to items "posted" (Adjusted R Square = .1536; p=.0008). This result is equivalent to a significant Pearson r. Items "read" was significant at p=001 and number of "hits" was significant at p=.0016.

The correlation matrix for the four variables including final grade was* **:

	Grade	Hits	Reads	Posts
Grade	1.00	.3869*	.4025**	.4118**
Hits	.3869*	1.00	.8081**	.6590**
Reads	.4025**	.8081**	1.00	.6670**
Posts	.4118**	.6590**	.6670**	1.00

*p<.05, **p<.01

These intercorrelations are somewhat higher than those obtained by Wang and Newlin (2000). This was probably due to the fact that the courses analyzed utilized the readings and postings of an online course to a greater extent than the psychology statistics sections studied by these researchers. Newlin (2000). Along this line of reasoning and taking into account the small number of cases for each section, separate section correlations are presented. The frequencies are higher by two for those correlations for American Literature-Colonial to Civil War and New Media Technology that do not include final grade.

English Composition (n=12)

	Grade	Hits	Reads	Posts
Grade	.1.00	.2540	.1288	.5372
Hits	.2540	1.00	.5148	.5125
Reads	.1288	.5148	1.00	.4769
Posts	.5372	.5125	.4769	1.00

Creative Writing (n=14)

	Grade	Hits	Reads	Posts
Grade	1.00	.7203**	.6578*	.8221**
Hits	.7203**	1.00	.9739**	.7867**
Reads	.6578*	.9739**	1.00	.7045**
Posts	.8221**	.7867**	.7045**	1.00

American Literature Civil War-20th Century (n=15)

	Grade	Hits	Reads	Posts
Grade	1.00	.3779	.2447	-.1035
Hits	.3779	1.00	.7196**	.1388
Reads	.2447	.7196**	1.00	.0537
Posts	-.1035	.1388	.0537	1.00

American Literature-Colonial Period-Civil War (n=13)

	Grade	Hits	Reads	Posts
Grade	1.00	.4058	.4645	.3679
Hits	.4058	1.00	.7314**	.4844
Reads	.4645	.7314**	1.00	.4922
Posts	.3679	.4844	.4922	1.0

New Media Technology (n=10)

	Grade	Hits	Reads	Posts
Grade	1.00	-.0226	-.2571	-.4626
Hits	-.0226	1.00	.3141	.5400
Reads	-.2571	.3141	1.00	.6268*
Posts	-.4626	.5400	.6268*	1.00

Each instructor was interviewed to see if the correlations could be explained for his/her course. In addition, other instructors of online courses were interviewed.

Case Interviews

Case Interview I – Professor of English Composition and Creative Writing:

For the English Composition class, there is only a minimum of instruction and evaluation that takes place interactively. Students mostly attach papers via e-mail and the papers are graded. The class discussion never gets "gelled." With English Composition, a mix of activities is required. The percentages for determining the final grade are seven essays worth 35 points, forum postings worth 30 points, private mail worth 25 points, final argument worth 5 points and a final vocabulary exam for 5 points. There is interactivity in English Composition but it is not as great as with the Creative Writing course.

Creative Writing is extremely suited for online instruction. The professor uses student papers and gets critiques from other students on the posting. The instructor regulates the submission of the postings and determines when to cut off discussion. She grades with check marks and +s. The final grade is still determined by both the quantity and more so by the quality of the postings. However, in general, students who have a greater number of "hits" and "posts" usually do better. They are more likely do the assignments. The students who do well are concerned about their writing and do excellent critiques. Because postings involve student writing and the more motivated students post a greater number of times, students with a greater number of posts tend to end up with higher final grades.

Case Interview II – Professor of American Literature

The American Literature professor taught two online courses that were used in the study. He was told about the direction and magnitude of the correlations that were obtained between final grade and number of "hits," number of items "read" and number of items "posted." He said that a significant correlation would be expected in his course, since the grade he gives is based on class discussion. Students must respond to information within the course and discussions are part of the grade.

He was skeptical about the number of hits as indicators of student learning. He explained that as students in regular face-to-face instructional settings can come to class and fall asleep, the same can happen for online courses. "Hits don't tell you what they've been learning." Students just increment the course counter that way.

This professor said that in his course students must comment on items and that the quality of postings is judged on one assignment per week, which is a short essay about a reading in literature. The instructor was convinced that participation in the course depended on what course is being taught. Therefore, the relationship between online activity and final grade depends on the course. With some courses the student can get online five or six times and it would be sufficient. With other courses, students need to log on a great deal more for interaction and participation.

Case Interview III – Professor of New Media Technology

The number of hits is the number of times the student entered the class. Postings are used when students post introductions. The number of postings is not as great in New Media Technology because private mail is used to screen what goes on to the bulletin board. When asked how a student with many hits gets a low grade, the instructor said it's probably the case that the student was insecure and couldn't get started. Also, with respect to postings and a low grade, the student may have asked for additional instruction via a posting. These postings include queries to the instructor. More than one posting may be submitted. In fact, if one were to examine the number of postings for New Media Technology, all but one had one posting for the introduction. The other student had two postings with a failing grade.

The combined correlation matrix after taking out the New Media Technology students improve the magnitude of the correlations as follows. The correlations are now based on 54 students.

	Grade	Hits	Reads	Posts
Grade	1.00	.4201 **	.4413 **	.4530**
Hits	.4201 **	1.00	.7554 **	.5810**
Reads	.4413 **	.7554 **	1.00	.5877**
Posts	.4530 **	.5810**	.5877 **	1.00

** p<.01

Interviews with Other Instructors of Online Courses

There were several short interviews with other instructors of online courses. One instructor who teaches accounting said that postings are rarely used since Excel spreadsheets are submitted and returned via e-mail attachments. In accounting, the course is less suited for online instruction. The same is true for a course in statistics. The instructor for statistics said that the "nature of the academic discipline is going to have some bearing on the nature and number of postings in the course" (Jones, 2002). In both the accounting and statistics courses, the instructors were trying to mirror the regular course in the online environment. This is in contrast to the creative writing course, which uses the online instruction tools as an enhancement. With the statistics course it is just a supplement. The students answer questions from their textbook and can send e-mail attachments to their instructor. In creative writing, their assignments and responses to the writing assignments of others are both posted. So number of posts will be more predictive of course grade when posts are a greater part of the course. However, in courses like accounting and statistics, the number of e-mails is counted and a record of is kept of this number.

Correlational Anomaly

Although with the combined group, significant correlations were obtained with final grade, there were students with very high numbers of "hits," "reads," and "posts" who did not receive a grade of A. A list of the top ten highest frequencies for each activity is presented below.

Top Ten Highest Frequencies for Each Activity and Corresponding Course Grade

No. of Hits	Grade	No. of Reads	Grade	No. of Posts	Grade
856	A	584	C	44	C
856	B	585	A	45	C
904	C	598	B	45	B
912	A	639	C	64	A
953	B	669	B	68	B
1071	B	703	B	75	C
1204	B	838	A	95	A
1205	C	864	B	111	A
1244	A	891	B	111	A
1513	A	896	A	123	A

Note the many B's and C's in the top ten numbers of "hits," numbers of "reads," and numbers of "posts." These data show a number of things. First, some students may log in many times without always doing work. Second, students may be logging in to retrieve content pages and do not read, comprehend or retain all the material. Third, the non-A grades associated with the highest number of "posts,"show that the qualitative assessment of the postings is more important than the quantity of postings.

Discussion

There are various issues that are beginning to emerge with respect to online courses. Some of these issues center on that part of online courses that substitutes for the lack of face-to-face instruction. Combining the data on the five course sections, correlations between final course grade and "hits," "reads," and "posts" were significant. Yet, computing the correlations separately by course section with smaller numbers told a different story. Courses in which the frequency of postings was a greater part of the course produced higher correlations with final grade. However, from the sets of correlations and from the case interviews with instructors, a number of issues emerged while another number of issues need to be given careful consideration.

First, there is the issue that different magnitudes of correlations will be obtained for different types of courses. For courses which use online activity like postings directly, there will be a higher correlation with course grade. Motivated students post more frequently. Fewer postings will be made in courses which use more emailing and which attempt to mirror the regular face-to-face instruction of the course.

Second, there is the issue of the quantity versus the quality of the postings. From an examination of the grades associated with the most frequent number of "hits," "reads," and "posts," it is apparent some students receive B's and C's. The reason the New Media Technology professor gave for this was that some students might be insecure. What she might have also meant was that students who login and post a great deal may be having difficulty with the technology and, especially if it's their first online course, may not know how to optimally use the technology. So they end up having a greater amount of online activity. And, in some courses, instructors count the frequency of posts as part of their participation grade.

A third issue related to the first is the amount of interactivity in the course as opposed to textual comprehension. Some courses are by nature suited for more interactivity. Other courses where the textbook is more important require less interactivity and fewer postings and more textual comprehension. In such courses one would expect lower correlations.

For the psychology statistics courses used in Wang and Newlin's (2000) study, moderate significant correlations were reported. They did not make conclusions across different disciplines. Here, separate sets of correlations are provided and they indicate different results.

Two policy questions raised by this study are whether online literacy be made part of placement tests and whether there should be mandatory training if students want to take an online course. Finally, different instructors of the same course may not produce the same quality course. Two instructors in the study did not use postings except for introductions. The question this raises is whether or not there should be faculty development in online course design.

Research has already been conducted comparing face-to-face instruction to online instruction controlling for student ability. However, that is beyond the scope of this paper. As online courses develop, new ideas will be tried and become more popular. One new idea is keeping track of the number of emails sent by students. However, it would not be radically different from what was presented here. One development which would be of interest to designers of software, like WebCT, would be to track the instructor's judgment of the quality of postings and look at the extent to which the qualitative ratings are related to grades in the course.

References

Finnegan, C. and Finnegan, B. (2002). Assessing the online classroom: Handy tools in three learning management systems. Paper presented at the 42nd Annual Forum for the Association of Institutional Research, Toronto, Ontario, Canada

Jones, J. (2002). Personal communication.

Kendall, M. (2001). Teaching online to campus-based students: The experience of using WebCT for the community information module at Manchester Metropolitan University. Education for Information, 19, 325-346.

Turoff, M. (1997). Alternative futures for distance learning: The force and the darkside. Invited keynote presentation at the UNESCO/OPEN UNIVERSITY International Colloquium, New York, NY.

Wang, A.Y. and Newlin, M.H. (2000). Characteristics of students who enroll and succeed in web-based psychology classes. Journal of Educational Psychology (92), 137-143.

USING GRID-GROUP THEORY TO UNDERSTAND STUDENTS AND INSTITUTIONS

Carol Trosset, Ph.D.
Director of Institutional Research
Grinnell College

Grinnell College[1] students have a reputation for being egalitarian, anti-business, and extremely politically liberal. Many of them do hold these views. Despite the fact that every year CIRP Freshman Survey results demonstrate the presence of those who want to succeed at their own businesses, make a lot of money, or think abortion should be illegal, no one thinks of these as "Grinnell" things to do or think. Furthermore, when CIRP data were subdivided into white students, students of color, and international students, it became clear that the extremely liberal student culture mainly characterizes Grinnell's white students. Students of color, and particularly international students, are far more likely than American white students to value leadership skills and socioeconomic status differences.

Grid-group theory provides a conceptual framework for describing and understanding these kinds of differences. Developed by anthropologist Mary Douglas[2], grid-group theory uses dimensions of social organization and ideology to compare cultures, institutions, and even individual proclivities on the basis of two independent dimensions. "Grid" measures the degree of restrictions placed on individual freedom of action. A high-grid society places limits on individuals by having many rules, a strict hierarchy, or some combination of types of regulation. "Group" measures the tendency of individuals to become incorporated into groups. High-group societies may show a tendency to incorporate outsiders, and/or a focus on us/them boundary maintenance.

Grid and group are unrelated dimensions of social organization. When used in combination to describe societies, four types of social organization result. Hierarchical societies are both high grid and high group, with a strong collective identity but significant internal differentiation, such as the English class system or the Indian caste system. Enclaves are high group but low grid, strong collectivities with few internal distinctions; many (though by no means all) small-scale tribal societies are of this type. Individualist societies, like the United States middle class, are characterized by low grid and low group. High grid and low group societies are known as Isolates; people who accept the presence of hierarchy and restrictions but who tend to focus their efforts on the circumstances of individuals or nuclear families rather than on whole communities.

---

[1] Grinnell College is a private, highly selective, liberal arts college in Iowa, with about 1400 undergraduates from all 50 states and many foreign countries. Of the student body, about 12% are students of color and about 10% are international students.

[2] Douglas, Mary. 1970. *Natural Symbols: Explorations in Cosmology*. Harmondsworth: Penguin. (and) 1978, "Cultural Bias" (Royal Anthropological Institute, Occasional Paper 35). London:RAI.

Applying Grid-Group Theory to CIRP Data

The Higher Education Research Institute's CIRP Freshman Survey includes many questions that measure aspects of either grid or group. For example, a life goal of having administrative responsibility would indicate high-grid tendencies, while wanting to promote racial understanding suggests high-group values. To operationalize grid-group theory in terms of CIRP data, I worked with a faculty colleague[3] to identify all CIRP questions that measured the theoretical components of either grid or group. Using unit-record Grinnell student data and SPSS software, I put all the selected items into an unrotated factor analysis. Applying a factor loading cut-off of 0.5, the first factor comprised nine grid items and the second nine group items. There were no correlations between the grid items and the group items.

The grid items identified were:

- want recognition from colleagues
- college's graduates go to top graduate and professional schools
- college's graduates get good jobs
- want administrative responsibility
- want to be well off financially
- want college to train student for a career
- want to be an authority in one's field
- want to be a community leader
- self-rated high drive to achieve

The group items identified were:

- expect to take part in protests during college
- expect to take part in community action
- want to help others in need
- want to promote racial understanding
- expect to do environmental clean-up
- want to influence social values
- participated in demonstrations in high school
- want to influence the political structure
- expect to do volunteer work

Using these 18 items, I calculated grid and group scores for each individual student, by summing the numeric scores assigned to their responses. Working with these individual scores, I ran correlations with all CIRP items not included in the factor analysis, to identify other variables that correlated highly with membership in one of the four social-organizational quadrants. (One caveat: it is not possible to locate precisely the dividing line between high and low grid or group. These are relative measures only, intended to be used in a comparative framework.)

---

[3] Anthropologist Douglas Caulkins, Grinnell's Earl D. Strong Professor of Social Studies.

ISOLATES competitive, good physical health, want to make money, politically conservative	HIERARCHISTS leadership skills, religious and political involvement, want to raise a family
INDIVIDUALISTS (no strong correlates)	ENCLAVISTS want to learn about interests, politically liberal, may drop out temporarily, smoke

I also calculated scores for each of a number of other selective private institutions. Grinnell's student body, similar to those of several other midwest colleges, appears to be relatively high group and low grid. Several northeastern colleges were somewhat higher grid and lower group, with two southeastern colleges still further in that direction. One private midwestern university had students with the highest grid tendencies. The regional clusters were, for me, unexpected, but they do conform to widely held perceptions of institutional cultures.

Grinnell, then, has a predominantly Enclavist student body. This conclusion has good face validity; there is a strong emphasis on "being Grinnell," and an egalitarian culture with few rules or requirements. Within the enclave, however, various groups of students differ from one another, based on their average grid and group scores:

- Asian-American, African-American, and international students have far higher grid scores than white students.
- Male students are slightly higher grid and lower group than females.
- Lower income students have much higher grid scores than high income students.

Grouping entering students by their intended majors also produces an interesting distribution, with pre-business and computer science students being the most Isolate (high grid, low group), and women's studies students being the most Enclavist (low grid, high group).

One advantage of the grid-group perspective is that it provides a different way to think about what "diversity" means. Ethnicity may remain a relevant dimension, but here the focus is on other variables highly relevant to higher education, such as career and life goals, and what kind of society students see as desirable. Grid-group analysis also shows clearly that not all students fit into the dominant institutional culture, and presents this in a visual way that may help make the point to a wide audience.

Cross-Validation: Faculty Views

The HERI Faculty Survey also contains a number of items that measure either grid or group. Some ask the faculty member to describe the institution, while others ask about the faculty member's own views and preferences. I followed a similar process of factor analysis (though with less strict cutoffs), and looked at the percent of the Grinnell faculty who agreed with particular items. Overall, the Grinnell faculty describes the college as Hierarchical, using the following items:

High Grid (Institutional Priorities):

- raising money
- enhancing national prestige and image

High Group

- faculty are committed to the welfare of the institution
- faculty are interested in students' problems
- administrators are interested in students' problems
- students and faculty interact outside of class
- students interact with each other outside of class

The following items were used to identify the preferences of individual faculty members:

Grid measures:

- want to be an authority in one's field
- want to obtain recognition from colleagues
- want to hold an administrative post
- want to be well-off financially

Group measures:

- want to be a good colleague
- want to provide services to the community
- want to promote racial understanding
- think the college should be involved in social problems

Using these items to calculate grid and group scores for individual faculty members, the Grinnell faculty can be divided as follows: 46% individualists (low grid, low group), 34% enclavists (low grid, high group), 11% isolates (high grid, low group), and 9% hierarchists (high grid, high group). Each group could be characterized by at least one related item. Hierarchists are the most dissatisfied with Grinnell's lack of curricular requirements (a low-grid feature of the college). Enclavists, as a group, place the highest value on gaining knowledge of other ethnic groups (an incorporative high-group value). Isolates are the most likely to find both colleagues and committees very stressful, while Individualists are the least stressed by faculty meetings. There were no significant differences between quadrants with respect to discipline, rank, gender, commitment to teaching, or overall job satisfaction.

Further Validation: Views on Leadership

In April 2002, as part of an apparently unrelated project (funded by the Lilly Foundation), I interviewed 28 Grinnell faculty members and asked them about student attitudes toward leadership. I told them that CIRP data consistently show Grinnell students as relatively unlikely to consider themselves leaders or to have leadership goals, and asked them why they thought this might be the case. Without knowing anything about grid-group theory, they gave me the following explanations:

- Students are skeptical of any power distinctions. (low grid)
- Students associate leadership with competition and hierarchy, and avoid it so as not to taint themselves with these qualities. (low grid)
- Students are not ambitious (low grid), and have a low regard for individual initiative to achieve things (high group).
- Students believe that if you can reach a consensus on something, it will just happen, without any need for a leader to do anything. (enclavist)

At the same time, the faculty are very aware of internal variation among the students in this respect, as the following comments illustrate:

- Students of color are more likely to see themselves as leaders, and have often participated in leadership programs at church or in high school.
- Foreign students are more ambitious, and see individual achievement as good.
- White students are the only group that sees leadership and achievement as linked to personal selfish agendas.
- Students who do understand leadership just grin and bear it here, waiting to graduate.

In addition, economics and chemistry professors were the least likely to have experienced anti-leadership bias on the part of their students.

Possible Uses for the Grid-Group Framework

The face validity of the results for both students and institutions suggest that grid-group theory provides a reasonable way to describe college culture. The appearance of themes related to grid and group in faculty interviews further suggests that the theory will make intuitive sense to members of the college. It may, therefore, be a useful way to educate people about internal cultural differences. Grid-group theory clearly can be helpful in examining the fit between students and institutions, since some entering students will fit the dominant local culture better than others. An imperfect fit might be seen as a problem (and might ultimately result in student attrition), or it might be seen as an opportunity to educate and acculturate the students in a different dimension of social adaptation.

Grid-group theory is purely descriptive and hence value-neutral. However, members of any particular quadrant may be inclined to criticize other types of organization. Grinnell College culture tends to be not simply low grid, but anti-grid. This poses a problem both for students who have higher-grid tendencies, and for the staff who are attempting to serve them. (Different types of institutions might pose similar problems students of a different description; for example, enclavist students might find themselves out of place at a business college.) Despite even a genuine commitment to serving diverse students, the kind of diversity described by grid-group analysis is one people do not think about much. Greater awareness might result in better efforts to accommodate different types, or in attempts to change the values of people who diverge from the local culture.

Several possibilities exist for expanding the analysis in this paper. More institutions of more different types could be compared, either as aggregates or as merged collections of individual students. First-year and senior survey data could be matched to determine whether students change in their persuasions over the four years. This analysis could be used to examine retention, testing the theory that students who fit the institutional culture would be more likely to remain and graduate. It would also be possible to identify direct measures of institutional culture that do not come from student or staff opinions. For example, the extent of curricular requirements would provide one measure of grid, while group could be defined in terms of parameters like the percent of students living on campus, or the percent of faculty who live nearby. Such measures would contribute to the investigation of both student-institution fit and of the effect of institutional type on students.

ASSESSMENT AND ASSISTING THE COLLEGE PRESIDENT STEER THE SHIP: AN ANALYTIC COMPARISON OF DASHBOARD INDICATORS, THE BALANCED SCORECARD, PERFORMANCE MEASURES, AND SIX SIGMA IN THE COLLEGE AND UNIVERSITY SETTING

Gail Wisan, Ph.D.
Director of Institutional Research
Goucher College
Baltimore, MD

According to a recent paper on excellence in the college and university setting by the Middle States Commission on Higher Education assessment has two purposes: accountability and effectiveness. This is true not only at the student learning level but also at the level of governance of the college or university. College presidents, Board of Trustees, Board of Regents, and legislators and governors all want data from colleges and universities for the purpose of accountability, to assess the quality of the college's performance. Regarding effectiveness, the focus is on measuring how successful the college is in producing results or outcomes.

There are many constituents and questions that drive the search for useful tools to help guide the college or university. Are the resources of the colleges and universities being used in the best way possible? What criteria can be used to assess the performance of a college or university? How can the Board of Trustees assess the performance of the college president and the executive vice presidents? How can a governor and/or legislature assess the performance of a college or university? How can a college president demonstrate her/his effectiveness and how can s/he best steer the ship? What tools are available to help the president steer the ship? Both accountability and performance tools provide feedback to the college's president and staff, encouraging improvement.

Many management tools have been used in corporate environments for purposes of honing performance. This paper reviews some of these management tools and examines their utility for the college or university setting, in general, and for the Board of Trustees and College President, in particular. Specifically, can these management tools, often from the corporate environment, be used in the educational environment to aid college presidents, Boards of Trustees, and governments in assessing the performance of colleges and universities?

Some of the tools that either have been used or are available to aid college presidents and other stakeholders include: dashboard indicators, key performance indicators, the Balanced Scorecards, the Baldrige Criteria for Performance Excellence Framework, Performance Measures, and Six Sigma. These management tools offer an array of choices to policymakers and accreditors.

This paper will: 1) provide a brief overview of five different management tools; 2) review the type of data and information used by these models; 3) organize the extant tools, providing information on how they compare in terms of providing helpful information and data for "steering the ship." *This paper will thus take a step towards organizing the available management tools and assessing their usefulness by providing a critical meta-assessment.*

Overview of Five Models and Management Tools

The five management tools that will be examined in this paper are: Dashboard Indicators, the Balanced Scorecard, Baldrige (the Baldrige Education Criteria for Performance Excellence Framework), Strategic Performance Measures, and Six Sigma. These tools and processes offer an array of choices as models for organizational assessment and management. While many are used in the business and corporate world, the paper will show how they can be useful in the educational, especially the nonprofit educational, world.

Dashboard Indicators

The Dashboard is a metaphor that draws from the idea of a dashboard, the panel under an automobile or airplane's windshield that contains various dials, gauges and indicators, instruments that register how the vehicle is performing. The indicators provide a warning system when the automobile or airplane is in danger. The indicators should be judged, paralleling the car or airplane dashboard, by their ease of use, importance, and clarity. If you have too many dashboard indicators, you may become distracted and miss the most important and basic ones. Thus, Dashboard Indicators tend to be one-page documents that draw on important information and which are presented in a succinct, easily understood, visually appealing manner. Conventionally, color is used to indicate good or bad performance, with green indicating good performance and red poor performance. Dashboards can be applied to many different policy areas but are usually intended for use by chief administrators, overseeing boards, and political decision-makers.

Balanced Scorecards

The Balanced Scorecard is a management tool that was developed by Robert Kaplan, a professor at Harvard Business School, and David P. Norton and presented in their book, *The Balanced Scorecard: Translating Strategy into Action*. Their approach was in response to what Kaplan felt was an overemphasis on the financial perspective when assessing organizational performance, which did not give a valid picture of that performance. Rather, he sought balance and added other perspectives: the customer, internal processes, and learning and growth. The Balanced Scorecard, taking into account the four perspectives, is a management tool that helps translate an organization's mission and vision statements into specific quantifiable goals and then monitors the organization's performance in terms of achieving the goals in these four areas.

Baldrige Performance Excellence Framework

The Baldrige National Quality Program, administered by National Institute of Standards and Technology, aims to improve the quality, productivity, and competitiveness of U.S. organizations in a variety of industries, from the private corporation to the nonprofit health and educational organization. Established in 1987 by Congress and named for the deceased Secretary of Commerce Malcolm Baldrige, the Baldrige Award was originally limited to for-profit companies that demonstrated excellence in performance. In 1998, NIST established new criteria for education and health care organizations and in 2001, educational organizations were awarded Baldrige Awards for the first time.

According to the Baldrige National Quality Program publication on "Educational Criteria for Performance Excellence," the criteria are intended "to help organizations use an integrated approach to organizational performance management that results in," 1)Improved value for students and stakeholders; 2) Improved overall institutional effectiveness, and 3) Improve organizational and personal learning." The Baldrige criteria cover seven key categories: 1) leadership; 2) strategic planning; 3) student, stakeholder and market focus; 4) information and analysis (i.e., fact based culture and decision making); 5) faculty and staff focus; 6) process management; and 7)organizational performance results. In looking at the weight of the above criteria, almost half the points are assigned to the seventh criterion--performance results, especially student learning. Other performance results include student and stakeholder-focused results (e.g., satisfaction), budget and financial results, market results, and faculty and staff results.

Performance Measures

In 1994, the Governmental Accounting Standards Board (GASB) issued Concepts paper number 2, "Service Efforts and Accomplishments Reporting." According to this, accountability requires performance measurement and, specifically, that government entities should report on results achieved with resources. GASB emphasizes the need for: 1) Clear goals and objectives; 2) set measurable targets to achieve; 3) develop quantitative indicators and then provide feedback on progress towards the objectives and goals. GASB associates performance measurement with Management for Results.

Six Sigma is a management strategy that began at Motorola when an executive argued that the company's problems were caused by its poor-quality, too- Six Sigma often-defective products. Six Sigma is a management strategy that argues that traditional quality control initiatives and programs were misguided, and guides companies to recreate processes so that defects are never occur in the first place. According to this strategy, improving quality is less expensive than poor quality, because the latter is very expensive to monitor. Moreover, correcting defects is very expensive, as is the loss of customer good will.

Based on the belief that argues that with higher quality come lower costs, not the opposite, Six Sigma has the goal of improving profit margins with the indirect results of improving quality. The focus is on analyzing process and on focusing on the processes that have the biggest impact on profit margin and customer satisfaction.

The Six Sigma strategy focuses on measuring defects per million opportunities (for defects). It doesn't matter what the product, process, or service is. When defects are eliminated, the time and money spent on detecting and fixing defects can be applied to improving the service or product.

Motorola focused on process improvement rather than detecting and fixing defects as the means to true quality improvement. In fact, the recreated processes so rarely produced defects that it was not cost effective to check for them. Six Sigma uses exact measurement to anticipate problem areas and stresses the importance of being proactive, rather than be reactive. Motorola launched Six Sigma in 1987 and two years later they were honored with the Malcolm Baldrige National Quality Award.

Very important in Six Sigma implementations is determining key processes, measurement, statistics, and "black belts." These black belts do not, contrary to some humorists, beat people into submission. Rather, they are project leaders with the task of leading key staff to be black belts who will lead staff through improving the key processes. Black belts play a key role in the implementation process.

The implementation steps for Six Sigma are:

1. Recognize: the process with the problems (Projects with the largest customer and revenue impacts are selected);
2. Define: goals, customer specifications; (D)
3. Measure: feedback; (M)
4. Analyze: specific statistical tools; (A)
5. Improve; (I)
6. Control: monitor. (C)

The term Six Sigma comes from the distribution in the Normal Curve. At the Six Sigma level, 99.9999996 percent of the distribution is included. Motorola say defects will only be two per billion opportunities. Sigma is, of course, the symbol (i.e., the Greek letter) used for the standard deviation. The significance of the concept and title The Six Sigma Way is that while focusing on "core processes and key customers," a company can reduce defects to practically zero, removing the need for and cost of quality correction programs. Six Sigma does not seek to improve quality for its own sake but rather, by using measurement and statistical tools, aims to improve processes, thus increasing profits. It seeks to analyze throughput defects, frequently left undetected, building upon the work of Deming who argued that defects are created at many points along the production process and that the causes of variation should be identified and fixed. Six Sigma emphasizes the role of measurement and statistical analysis for identifying the causes of variation and the use of statistical tools (e.g., graphs) for providing data to eliminate the variation (i.e., the defect).

Comparative Critical Analysis of the Management Tools

Clarity/Ease of Comprehension

A management tool that is succinct and quickly grasped is superior to one that requires review of many pages of a document. The parallel of the dashboard was used to indicate the desire for a small set of key measures that could provide the organization's most important information and allow the president or board of trustees to quickly assess whether the organization is performing at a superior, average, or inferior level. The two tools that perform best on this quick, clear, and digestible scale are the dashboard and the Balanced Scorecard. Performance measures are also good in this area, but they are either not as visually appealing or they may not provide enough measures. More so with performance measures than dashboards and balanced scorecards, it may depend upon the specific implementation of the model and tool. Baldrige criteria and Six Sigma are not tools that emphasize a clear concise communication methodology.

Indicator Selection Guidance

The dashboard does not usually specify particular administrative areas on which to focus; it is more suggestive of a mode of delivery than of a substantive area. Nonetheless, in practice, the lack of specific guidance regarding topics to measure means that standard measures tend to populate the dashboard, regardless of whether they are actually the key ones needed to gauge the performance of the organization. One of the strengths of the Balanced Scorecard is that it does direct the organization to focus on not only financial measures, but also measures related to customers, internal processes, and learning and growth. While the Balanced scorecard may not identify the specific indicator or measure to be used, it moves the organization in the right direction by specifying four essential areas to examine and providing a cascading methodology for moving from the top down to the unit level. The Cascading Balanced Scorecard moves from the high level mission down to the unit with specific measurable goals, measures with quantifiable indicators, targets and initiatives.

The Baldrige Framework provides specific criteria that help organizations measure and report on appropriate data. Thus, it provides guidance on the topics to include and even demonstrates how thoroughly the Baldrige awards assess the specific criteria by attaching a number of points. This provides real guidance to the organization that uses these criteria for self-assessment and improvement.

Performance measures like the Dashboard measures are vague in their application. What this can mean, in practice, is that standard measures (e.g., FTE, graduation rate, and percentage of minority groups) are used. As with traditional accountability indicators, inclusion is based heavily on traditional educational indicators and topics of concern, with little attempt to order them in terms of significance to the organization for either mission or performance. In the recent past, Performance Measures have been linked to the Management for Results movement. Thus, governmental organizations' self-assessment and performance measures are frequently set in the context of MFR, or

Management for Results. In Maryland, public colleges and universities and even the Higher Education Commission itself are expected to report their performance measures in the context of MFR. MFR provides some guidance in the choice of indicators. The indicators are linked, theoretically but not always in practice, to the specific college or university's mission, a positive step in the assessment process. The negative side of this approach is that in order to keep reporting brief, the number of indicators may be very limited and thus not provide a full picture. The political context of this struggle in Maryland led to a compromise between Maryland Higher Education Commission's list of Accountability Indicators and the new Management for Results approach.

The Six Sigma management tool does not focus on any required content for assessment. Rather, it focuses on measuring the performance of core processes, especially those that have the greatest impact on revenue and key customers.

Types of Data and their Significance

In helping the college president steer the ship, certain data-- historical data and comparative-- are very important. While some of the management tools use and report such data systematically, others do not.

Baseline Data

Baseline data are often provided implicitly in the form of current or the last data that is available. The sample Dashboard used in this paper (This can be obtained from the author) implicitly uses baseline data (from the previous year or last report); it then uses a color code to indicate better, worse, or the same performance since then. Baseline data are not specifically shown since the reported numbers are current, high and low for the past five years. The indicator "measure" in the Balanced Scorecard provides the baseline data from which to evaluate and develop the target.
All the Management Assessment models provide baseline data.

Historical Data

Historical data provide information that can be used for trend analysis. Compared to the past, the *trend* of the performance on the indicator can be tracked and monitored. Historical data are related to baseline data but are not exactly the same. The use of historical trend data is implicitly grounded in the quasi-experimental designs described by Campbell and Stanley in their classic work on research design in the real world. Trend data, multiple observations across time, permit assessment of whether a trend already existed prior to any policy or is a result of a change implemented by the organization. For example, enrollment might have been increasing for several years prior to a process change in the Admissions' Office. Without the historical trend data, it would be difficult to analyze how much (if any) of the increase in applications was a result of the internal process change and how much was a result of other factors that already existed (e.g., a demographic bulge).

Comparative/Peer Data

Besides historical data, the president and board need a comparative framework to assess organizational performance. For example, trend data for one college may show that its return on investment is down. But we might find that the return on investment for other comparable colleges for the same time period are down more and thus the college is actually performing better than its peers. Trend data for a higher education institution and its peers provide a framework to evaluate the performance of the higher education institution. Going one step further, Campbell and Stanley's classic book on quasi-experimental designs sheds more light on the value and peril of using comparative data without historical or, at least, baseline data. Without baseline or trend data for the comparison group, we do not know if the institutions are comparable. For example, suppose a retention program is instituted at a school and the results show an improvement in retention. Suppose two nearby schools are selected to study the effect on retention of a new program that aims to improve academic support. The program is instituted at only one of the schools where retention rates increase. At the other college that did not institute the program, retention rates are measured and are found to be higher than at the school that instituted the improvement in the academic support process. The better rates of retention at the other college do not provide evidence that the process improvement failed. Rather, the college without the process change may be a non-comparable control; it may have had higher retention rates before. (If we had measures of retention at both colleges before and after the process change, we could compare the rate of change in the two schools.) Peer data are important, but it is important that the "peer" data actually be comparable if we are using them to evaluate a college's performance.

In the Dashboard example used in this paper, trend data are used but peer data are not provided in the framework. Some dashboards provide peer data, but few trend data. The Balanced Scorecard provides baseline data, but does not highlight trend or comparative data. The Baldrige criteria involve trend data and comparative data to demonstrate performance excellence but these are not clearly highlighted. Performance measures and Management for Results reporting tools include historical data, but not necessarily peer data. Some Management for Results frameworks, such as Maryland's MFR, include peer data.

Six Sigma distinguishes between data from a "benchmark" institution and peer data. Six Sigma proponents argue that organizational leaders should use as a model an organization that is performing best, regardless of whether it is their industry. They argue, that if a college's peers are performing poorly, one should not benchmark a process against these, but instead should reach outside of the education industry for successful models.

Linkage to Mission or Strategic Plan

Only some of the management frameworks tie their tools to the organization's mission, vision, or strategic plan. Management tools that are linked to a college's vision or strategic plan offer a definite advantage. The Balanced Scorecard, the Baldrige

Criteria, and sometimes Performance Indicators are linked to the vision, mission or strategic plan. It is hard to reach a goal if leadership is not monitoring indicators that will permit them to assess whether the college is making progress.

Communication Tools for the Organization

The Dashboard is a clear and brief tool that facilitates communication, but it does not specifically include a communication strategy. The Balanced Scorecard is a superior tool for organizational communication because the "cascading Balanced Scorecard" requires that the mission and measures reach down to the unit level and that they be tied to the college's mission. Six Sigma uses "black belts" to help carry the message to the staff working on a specific project.

Comparing Management Assessment Tools On Various Dimensions

	Dashboard	Balanced Scorecard	Baldrige Criteria	Performance Measures	Six Sigma
Brief and Quickly Grasped	Yes	Yes	No	Yes	No
Directs Indicator Selection	No	Yes	Yes	No	No
Trend or Baseline Data	Yes	Yes	Yes	Yes	Yes
Peer Data	Some	No	Some	Some	Some
Explicitly Tied to Mission	No	Yes	Yes	Yes	No
Explicit Organization Communication	No	Yes	No	No	Some
Innovation	No	Yes	Yes	No	No

Innovation

The Balanced Scorecard includes innovation and growth among its four factors for measuring organizational effectiveness. Neither the Dashboard nor Performance

Measures include this as an explicit indicator. Among the Baldrige Awards' hoped-for outcomes are improved organizational and individual learning. Moreover, since one of the Baldrige educational criteria is strategic planning, one could argue that the Baldrige framework includes innovation and growth as dimensions. The Dashboard, Performance Measures, and Six Sigma do not focus on innovation. Six Sigma does sometimes lead to innovation by recreating processes; however, its focus is on improving internal processes and not on issues such as innovating for improved market share, e.g., by changing the product instead of just the process. This issue speaks to at least some educational institutions.

Discussion and Conclusion: Positives, Negatives, and Limits

On the positive side, a Dashboard provides a succinct amount of information in a limited space. The Dashboard and the Balance Scorecard are useful tools to provide a president and board of trustees with data on an institution's performance. On the negative side, some tools, for instance the Baldrige criteria, may be too cumbersome to use as a regular management tool. The Dashboard by itself does not provide adequate direction in the selection of indicators. Regarding limits, these tools require a serious commitment from the top down to be useful. If they are used simply for display rather than improvement or change, the tools will have little impact. All of these tools require valid data, a desire to truly examine the information, and a commitment to apply resources to priorities. Without the desire to actually confront the data and apply resources to implement specific key projects, these tools may end up being of limited utility. Moreover, it is not easy to overcome organizational inertia. Recognizing this problem, Six Sigma proponents have developed the concept and role of the "black belt" to help select a key project to implement and then lead it to fruition.

Example of One College's Dashboard

In Appendix A, the current (but under revision) Goucher College Dashboard Indicator is attached. This model offers a number of advantages:

Five years of data are analyzed;
The highest value for a five-year period is displayed;
The lowest value for a five-year period is displayed;
The current data value for an indicator is displayed;
There is a color-coded key that indicates the performance level (better is blue; worse is red; black is neutral).
The direction of change is indicated graphically: higher: >; lower: <; no change: 0.

A disadvantage of the Dashboard example is that peer data is not included for comparison and benchmarking. The current Dashboard includes items such as study abroad, that are directly relevant to the college's strategic plan but the linkage is not as explicit as with some other management tools, such as MFR. A shortcoming of the current Dashboard that will be changed on its next incarnation is the inclusion of customer satisfaction data. Items from the National Survey of Student Engagement and the Alumni Survey will probably be added.

References

Baldrige National Quality Program. (2002) "Education Criteria for Performance Excellence." Gaithersburg, MD. National Institute of Standards and Technology.

Baldrige National Quality Program (2002) "Baldrige, Six Sigma, and ISO: Understanding Your Options." Gaithersburg, MD. National Institute of Standards and Technology.

Berkman, Eric. (2002) "How to Use the Balanced Scorecard." *CIO Magazine.* (May 15), pp. 1-7

Borden, Victor and Banta, Trudy.(ed.) (1994) *Using Performance Indicators to Guide Strategic Decision Making*. (Summer: 92). New Directions for Institutional Research number 82.

Breyfogle III, Forrest W. (1999) Implementing *Six Sigma: Smarter Solutions Using Statistical Methods.* New York: John Wiley & Sons.

Burke, Joseph. (2002) "Accountability for Results--Ready or Not." *Trusteeship. (10:1), pp. 8-13.*

Campbell, Donald and Julian Stanley. *(1963) Experimental and Quasi-experimental Designs for Research.* Chicago: Rand McNally & Co.

Corts, Thomas E. and James C. Eck. (2002). "Ten Ways to Track Performance." *Trusteeship* (10:1), pp. 14-18.

Government Accounting Standards Board (GASB). (1994) "Service Efforts and Accomplishments." Summary of Concept Statement No. 2. http://accounting.rutgers.edu/gasb/st/concepts/gconsum2.html

Kaplan, Robert and David P. Norton. *(1996) The Balanced Scorecard: Translating Strategy into Action.* Boston: Harvard Business School Press.

Pande, Peter, Neuman, Robert, Cavanagh, Roland. *(2000) The Six Sigma Way: How GE, Motorola and Other Top Companies are honing their Performance.* N.Y: McGraw-Hill.

Ruben, Brent. (1999) "Toward a Balanced Scorecard for Higher Education: Rethinking the College and University Indicator Framework". *Higher Education Forum.* (Fall: 02), pp. 1-10.

The National Center for Public Policy and Higher Education. "*Measuring Up 2000: The State by State Report Card for Higher Education.*" (2000) http://measuring up2000.higher education.org

CLOSING THE ASSESSMENT LOOP: APPLYING RESULTS OF A PRIMARY TRAIT ANALYSIS TO IMPROVE EDUCATIONAL OUTCOMES

David W. Wright
Associate Professor of Management
West Liberty State College

Elizabeth A. Robinson
Interim Dean, School of Business Administration
West Liberty State College

Marsha V. Krotseng
Chief Planning Officer
Valdosta State University

Purpose

This paper describes the development, implementation and results of a student outcomes assessment program in a School of Business Administration (SBA) at a public baccalaureate institution. Specifically, we will discuss the results of a Primary Trait Analysis instrument and other assessment techniques (including student satisfaction surveys, focus groups, standardized testing) leading to any curricular, instructional and advisement changes within the SBA. The measures that were established and monitored through this process have provided valuable feedback for improving both school and institutional performance. This paper will focus primarily on the conclusions and changes resulting from the assessment process, hence, closing the assessment loop. Several unique elements of this assessment effort are that it was designed during the Fall 1999 semester and implemented in Spring 2000 -- a very aggressive timeframe; that it was accomplished with the active support and participation of all SBA faculty and administration; and that it was implemented at no additional cost to the School or the institution. Although evaluation the results is continuing, the process has produced useful information based on student, faculty and external review feedback.

Background

The School of Business Administration is one of four Schools within this state-assisted baccalaureate degree granting institution of approximately 2,600 students. The School comprises the departments of Administrative Systems and Financial Systems. There are approximately 550 majors in the School of Business Administration and fifteen full-time faculty members. The SBA offers one degree, the B.S. in Business Administration. Within this degree program, students may choose from twelve specializations including management, marketing, accounting, computers, and economics. Each specialization requires students to complete approximately thirty course hours in addition to forty-eight hours in the required business core. The business core includes instruction in management, marketing, accounting, economics, finance, communications, legal environment, and computers. It is the business core that was evaluated using a Primary Trait Analysis instrument.

Literature

In Assessment Essentials (1999), Catherine Palomba and Trudy Banta describe Primary Trait Analysis (PTA) as one of many assessment techniques that can be useful for classroom as well as program assessment. The PTA identifies the key factors or traits that are used in evaluating an assignment or project, and a standard three- to five-point scoring scale is developed for each trait. Each score "is accompanied by an explicit statement that describes performance at that level" (p 164). The higher the score, the more clear, complete, and accurate is the student's performance on that particular trait. Specific examples of some of the Primary Traits that emerged through this process and their descriptions are provided below.

Methodology

The College has a standing assessment committee that meets on a regular basis. The committee comprises faculty representatives from each of the four Schools as well as the Provost, the Vice-Provost/Director of Institutional Research, a Dean, a Department Chair, and the Dean of Student Affairs. The College administration and the committee have been strong advocates for the assessment process as evidenced by their support for sending committee members to national conferences, holding on-campus seminars, and encouraging the use of external assessment consultants as appropriate. The SBA has established its own assessment committee that works in conjunction with this College-wide committee composed of faculty from both departments as well as both Department Chairs. The SBA representative to the College assessment committee serves as an ex-officio member to the SBA assessment committee. This committee meets on a regular basis and reports to the faculty and administration of the School; all recommendations from the committee require approval by the School's entire faculty and administration.

The first step in the assessment process was to determine the education outcomes expected by the School of Business Administration. The following outcomes were outlined based on institutional mission:

1. Students will develop critical thinking, decision making and problem solving skills in the application of appropriate business principles and practices.
2. Students will be proficient in computer applications.
3. Students will demonstrate verbal and written communication skills.
4. Students will be satisfied with the quality of curriculum, instruction, advisement, technology/facilities and extra-curricular activities within the School of Business Administration.
5. Students will meet entry level requirements for employment in business.

After gathering assessment process information through such means as conference attendance and the use of external consultants, the School decided to use Primary Trait Analysis as one mechanism to measure the desired outcomes.

Historically, all business majors are required to take the capstone course "Administrative Policies." This course offers an opportunity for all students to exhibit the knowledge that they have acquired during their matriculation, specifically emphasizing knowledge related to the business core. This course provided an appropriate and logical venue in which to begin to measure the educational outcomes of SBA majors.

Data Sources

The "Administrative Policies" course requires students to complete a comprehensive case analysis within a group/team setting. A formal presentation is made by the team to other students and faculty from the School. In order to identify the traits that would be assessed, all faculty attended the students' presentations during Fall 1999. After observing these presentations, individual faculty members developed lists of potential primary traits. Early in the Spring Semester 2000, the SBA assessment committee considered this information and compiled a working document of primary traits that could be used to assess student outcomes. After careful discussion, the committee agreed that the following six primary traits reflect the outcomes expected of all business majors based on material presented in the business core:

1. Critical Thinking
2. Accounting and Finance Knowledge
3. Marketing Knowledge
4. Use of Visual Aids
5. Oral Presentation Skills
6. Written Communication Skills

These six traits were presented to the faculty of the SBA, and they were approved unanimously. The traits were reviewed and approved by the School's external Advisory Council comprising representatives of local and regional businesses who provide feedback to the School. The next step was the development of statements specifying the exact outcomes for each trait that would correspond with each of the five levels on the evaluation scale. This represented one of the most time-consuming elements of the process, since a number of meetings were required before the faculty were comfortable that the statements enabled them to satisfactorily distinguish various levels of performance. For example, the statements corresponding with the five levels of performance in Critical Thinking were:

5. Students exhibited an *advanced understanding* of Business Principles by *interpreting information, using appropriate models and techniques* (financial ratios, strategic management matrices, economic concepts, etc.) and were able to *logically draw conclusions* and *make appropriate strategic recommendations*. In addition, students were *able to defend their recommendations*.

4. Students exhibited an *advanced understanding* of Business Principles by *interpreting information, using appropriate models and techniques* (financial ratios, strategic management matrices, economic concepts, etc.) and were able to *logically draw conclusions* and *make appropriate strategic recommendations*. Students *had difficulty in effectively defending their recommendations.*

3. Students exhibited an *understanding* of Business Principles by *interpreting information, using appropriate models and techniques* (financial ratios, strategic management matrices, economic concepts, etc.). Students were able to *draw conclusions (not necessarily logical) and make appropriate strategic recommendations*. Students were *unable to defend their recommendations.*

2. Students exhibited *some understanding* of Business Principles but *did not properly interpret information or apply business models and techniques* (financial ratios, strategic management matrices, economic concepts, etc.). Students *failed to draw conclusions or make strategic recommendations*.

1. Students exhibited *no understanding* of Business Principles. Students *failed to interpret information or apply business models and techniques* (financial ratios, strategic management matrices, economic concepts, etc.). Students *failed to draw conclusions or make strategic recommendations*. *

*These descriptors were last modified in 3/19/01.

Additional data sources are as follows:

- CIS 471 - Advanced Computer Applications required of all senior business students where a comprehensive final is administered to evaluate proficiency in computer application skills.
- GBUS 490 - Professional Development Seminar (1 hr.) required of all senior business students where a Graduating Student Survey (including 20 specific statements about the SBA) is completed by the students along with students participating in a focus group designed to solicit more detailed explanation about perceptions of the SBA.

Results

Since April 2000, all SBA faculty have participated each semester in visiting the capstone course classes and evaluating the students' team presentations using such specific statements as those presented above. Two to three faculty members attend each presentation on a rotating basis. Throughout this time, the SBA has modified the trait descriptors based on faculty recommendations. In addition, the validity of the evaluations has been monitored by examining the level of inter-rater validity.

A sample of the assessment data collected is as follows:

	Spring 2000*	Fall 2001*
Critical Thinking	76.5	78.6
Accounting/Finance	60.8	78.6
Marketing	74.5	85.7
Visual Aids	56.9	91.7
Oral Presentation Skills	58.9	92.9
Written Communication Skills	60.8	71.4

*Percent of student groups scoring 3 or higher in faculty evaluations

Since the collection of the Spring 2000 data, the following changes were made:

- Descriptors of Primary Traits have been revised based on faculty feedback.
- Faculty workshops have been conducted to enhance the SBA faculty's ability to evaluate student performance in capstone course, Administrative Policies.
- Instructor in the capstone course clarified the instructions and requirements for the capstone project by making the instructions more descriptive for better student understanding.
- Instructors in the Principles of Accounting classes have placed a greater emphasis on the analytical approach to problem solving; student projects include analyzing corporate financial reports.
- Curriculum change: The course description for GBUS 322 - Business Communications was revised to require individual oral presentations using presentation management software.
- Writing Across the Curriculum was adopted in some business classes.

In the senior level computer applications class, a comprehensive exam has been administered in two consecutive semesters. Results are inconclusive as Computer

Information Systems faculty believe that the exam may be too elementary. Therefore, a new exam is being developed.

Results from the Graduating Student Survey administered in the Professional Development Seminar each semester have consistently shown that at least 75% of the students surveyed "Strongly Agreed" or "Agreed" with 18 of the 20 positive statements about the SBA—deficiencies in Advisement and Extra-Curricular Activities were noted. A questionnaire covering these two topics has been developed to better define these deficiencies and will be handed out this fall after pre-registration.

In the focus groups during the same period of time, students expressed a strong sense of satisfaction with the following areas in the SBA: Quality of Instruction, Student/Faculty Interaction, Curriculum, Advisement, SBA Organizations, Technology and Physical Environment of Classroom. As a result of student comments in the focus groups, these are some of the changes that have been implemented: SBA curriculum advising sheets now indicate the semester in which specialization courses are offered; a glass-encased bulletin board is being used to communicate with students about SBA organizations, activities and upcoming events; and computer labs now have extended hours to better serve our students.

Conclusions and Implications

This case study suggests that a sound assessment technique can be identified and implemented within a short period of time (one semester) and sustained given willingness, enthusiasm, and commitment by the School committee and faculty. The committee, with the approval of SBA faculty, has continued to make revisions to the process and propose necessary changes to curriculum, instruction and advisement.

Additionally, the assessment process was reviewed by two external consulting bodies. An internal review of the SBA (2001-02 AY) was conducted by an external consultant, Dr. Andrew Sikula, Sr., Director of Marshall University Graduate School of Management. He summarized the SBA assessment as "exemplary." The National Center for Higher Education Management Systems (NCHEMS) was contracted by the state of West Virginia to review all the assessment efforts in the state college systems. The first suggested recommendation made by these consultants was that West Liberty State College should "consider using Business Administration as an example, or case study, in the NCA [February, 2003] report."

301/405-8134

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC
Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

TM034885

I. DOCUMENT IDENTIFICATION:

Title: North East Association for Institutional Research 29th Annual Conference Proceedings

Author(s):

Corporate Source:

Publication Date: March 2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Beth A. Simpson, Membership Secretary

Printed Name/Position/Title: Beth A. Simpson, Membership Secretary

Organization/Address:
NEAIR
550 W. Frederick St, Lancaster PA 17603

Telephone: 717/358-7159

FAX: 717/358-4456

E-Mail Address: neair@fandm.edu

Date: 4/29/03

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Assessment and Evaluation
University of Maryland, College Park
1129 Shriver Lab
College Park, MD 20742

EFF-088 (Rev. 4/2003)-TM-04-03-2003